Impact of varying metal ion- and carbohydrate concentrations on gene expression in the human pathogen Streptococcus pneumoniae

Shafeeq, Sulman

Impact of varying metal ion- and carbohydrate concentrations on gene expression in the human pathogen *Streptococcus pneumoniae*

Sulman Shafeeq

The work described in this thesis was carried out in the Molecular Genetics Group (MolGen) of the Groningen Biomolecular Sciences and Biotechnology Institute (Faculty of Mathematics and Natural Sciences, University of Groningen, The Netherlands) and was financially supported by Higher Education Commission (HEC), Government of Pakistan.

university of groningen

Printing of this thesis was financially supported by the Groningen Biomolecular Sciences and Biotechnology Institute (GBB) and University of Groningen.

Cover design:	The cover shows a Scanning Electron Micrograph (SEM) of *S. pneumoniae* taken from Public health image library (image number 263). The cover was designed with the help of Laetitia Attaiech.
Printed by:	Ipskamp B.V.
ISBN:	978-90-367-5766-9 (Printed version) 978-90-367-5767-6 (Digital version)

RIJKSUNIVERSITEIT GRONINGEN

Impact of varying metal ion- and carbohydrate concentrations on gene expression in the human pathogen *Streptococcus pneumoniae*

Proefschrift

ter verkrijging van het doctoraat in de
Wiskunde en Natuurwetenschappen
aan de Rijksuniversiteit Groningen
op gezag van de
Rector Magnificus, dr. E. Sterken,
in het openbaar te verdedigen op
vrijdag 26 oktober 2012
om 14:30 uur

door

Sulman Shafeeq

geboren op 13 mei 1984
te Multan, Pakistan

Promotor:	Prof. dr. O.P. Kuipers
Copromotor:	Dr. T.G. Kloosterman

Beoordelingscommissie:	Prof. dr. J.M. van Dijl
	Prof. dr. ir. J.D. van Elsas
	Prof. dr. P.W.M. Hermans

Dedicated to my beloved parents

Abstract

Streptococcus pneumoniae is a Gram-positive human pathogen that is responsible for millions of deaths each year especially in children and the elderly, due to causing infections like pneumonia, meningitis, otitis media and sepsis. Little is known about the interaction of this human pathogen with and its behavior in the fluctuating nutritional environment inside the human body. This thesis describes several gene regulatory responses of *S. pneumoniae* to varying metal ion- and carbohydrate availability, which it is likely to encounter in different niches inside the human body during the infection process. Several transcriptional regulators responsive to various metals (*e.g.* zinc and copper) and carbohydrate sources (*e.g*. cellobiose) are characterized and their regulons studied in detail. The results described in this thesis will help to better understand the molecular biology of *S. pneumoniae* by providing new insights into the responses of *S. pneumoniae* to changing environmental conditions.

Contents

Chapter 1

General introduction

Streptococcus pneumoniae: A human pathogen

Bacteria (prokaryotes) are the largest group of single-celled living organisms on earth. There are approximately five nonillion (5×10^{30}) bacteria on earth (266). They have the ability to make complex symbiotic associations with other organisms and to survive in every possible niche on earth, including extremely harsh conditions, like deep in the earth crusts and in radioactive waste. These symbiotic associations can be divided into different lifestyles called parasitism, mutualism or commensalism. Bacteria that have the ability to interact with and inhabit the human body by symbiotic associations and that cause infections are called human pathogens. There are many human pathogens including several species in the *Streptococcus* genus that harm the proper functioning of the human body. The genus of *Streptococcus* consists of spherical Gram-positive bacteria belonging to the phylum **Firmicutes** and the order **Lactobacillales** (216). Several members of *Streptococcus* are important human and animal pathogens *e.g. Streptococcus agalactiae*, *Streptococcus mutans*, *Streptococcus pneumoniae*, *Streptococcus pyogenes*, *Streptococcus suis*, and *Streptococcus uberis*.

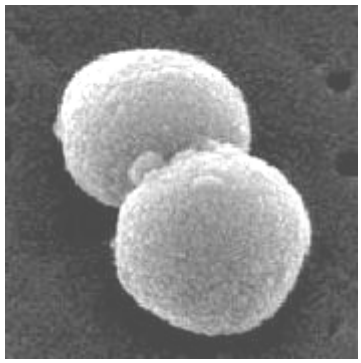


Figure 1: SEM (Scanning electron micrograph) of *S. pneumoniae* from Public health image library (image number 263).

S. pneumoniae (also called pneumococcus: Fig. 1) was firstly discovered by Stemberg and Pasteur in 1880 (14). It is an aerotolerant Gram-positive opportunistic human pathogen that has the ability to colonize the human nasopharynx as a commensal (31, 191). Pnemococcus usually stays in the nasopharynx asymptomatically until an opportunity arises that allows it to spread to sterile sites of the human body to cause disease (130). It generally colonizes the nasopharynx of the majority of children during their early childhood and sometimes a child/person can get colonized with more than one pneumococcal strain (42). During colonization, *S. pneumoniae* not only competes with the human immune system but also with other respiratory human pathogens including *Staphylococcus aureus* and *Haemophilus influenzae* (31, 158, 191, 226). To overcome the co-inhabitant competition for ecological resources including the availability of nutrients for colonization, *S. pneumoniae*

secretes different proteins among which antimicrobial peptides (Bacteriocins) (59). Previously, it has been described that bacteriocins, pili, hydrogen peroxide and various other factors play a major role in the competition for nasal colonization (59, 158, 204, 205).

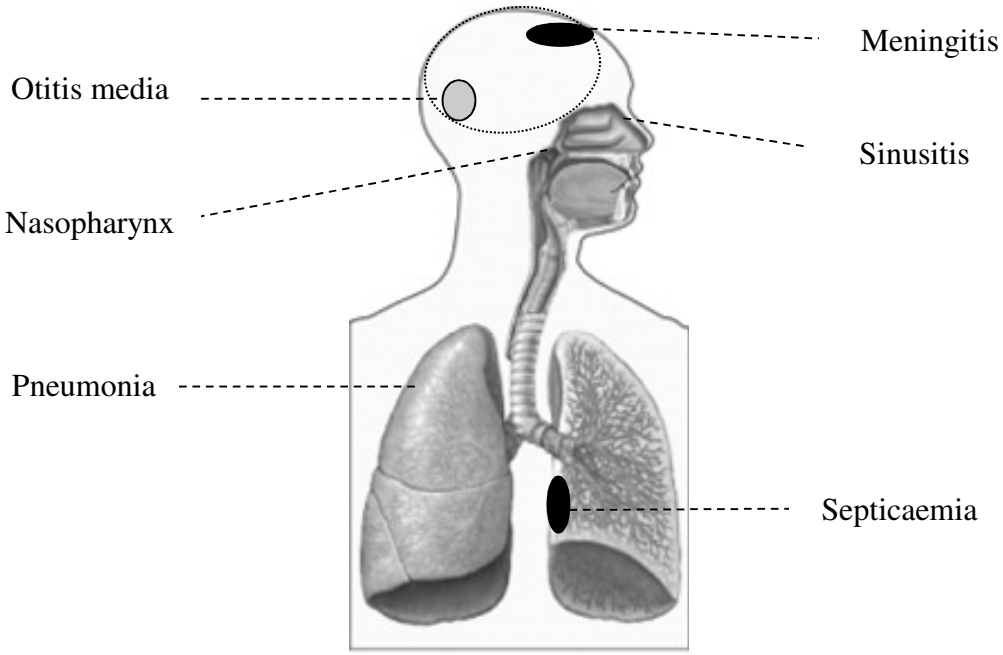


Figure 2: Diseases that might be caused by *S. pneumoniae* in the human body (www.sciencedaily.com).

S. pneumoniae is considered as a main causal organism of pneumonia, bacterial meningitis, bacteremia and otitis media (Fig. 2) (182). It more readily causes infections in young children and elderly people with a weak immune system as compared to adults with healthy immune systems. It is responsible for more than one million deaths especially in children each year worldwide (113, 182). Most of the infected children that die due to pneumonia are under the age of five (221). The morbidity and mortality rate due to pneumococcal diseases are high in both developing and developed countries. The increased pneumococcal resistance to drugs and vaccines over the passage of time has led to the evolution of more virulent serotypes of *S. pneumoniae*, which is also a reason for increases in the mortality rate caused by this bacterium. It is important to understand the regulatory mechanisms and roles/functioning of virulence genes in *S. pneumoniae* to find the reasons for pathogenesis and drug resistance, and to identify possible points of intervention for treatment or vaccination (171).

The introduction of this thesis highlights several points concerning the molecular biology of pneumococcus, including the role of virulence factors and the role of environmental changes like varying metal ion- and carbohydrate availability on gene regulation of *S. pneumoniae*. Findings in this thesis support the idea that proper gene

regulation tuned to the changing environment is of pivotal importance for the fitness and virulence of *S. pneumoniae*. The first part of this thesis shows the importance of metal ions in virulence and gene regulation, while the second part deals with the ability of *S. pneumoniae* to respond to non-preferred carbon sources and its effect on gene regulation.

Virulence factors of *S. pneumoniae*

Several genes are known to be involved in the virulence of *S. pneumoniae* and many of them are already characterized with respect to their involvement in virulence. One of the most important virulence determinants in *S. pneumoniae* is its capsule polysaccharide (CPS), which is surrounding the cell wall (201, 262). The capsule of *S. pneumoniae* does not play an important role in colonization as it impairs efficient adhesion to human respiratory epithelial cells (3, 58, 264). However, it protects pneumococcus against phagocytosis and helps in systemic dissemination after invasion (39, 147).

Another 'classical' pneumococcal virulence factor that was studied intensively over the last decades is pneumolysin (Ply). Ply is reported to exist in more than 20 different species of Gram-positive bacteria, including *S. pneumoniae* (171). Ply is conserved in almost all the clinical isolates of *S. pneumoniae* and expressed in the late log phase of growth (24). Earlier it was thought that Ply is expressed during the autolysis of the cell (25), but a recent study has shown that it can also be released independently of autolysins (19). It is quite important for the pathogenesis of pneumococcus, as strains lacking Ply are less virulent(41, 114). However, it plays no role in the colonization by *S. pneumoniae* (215).

Choline-Binding Proteins (CBPs) are a group of proteins that actively participate in the virulence of *S. pneumoniae*. CBPs are cell surface anchored proteins that are linked to the cell wall via their choline domains. The role of many CBPs like LytA (autolysin), LytB, LytC, CbpE, PspA (Pneumococcal Surface Protein A) and PspC (Pneumococcal Surface Protein C) has already been studied in *S. pneumoniae* and their role in virulence was established (76, 78, 143, 193, 198). LytA is one of the most important cell wall hydrolytic enzymes and plays a role in cell wall degradation, and release of DNA and cytoplasmic proteins like Ply (27, 171). Other cell wall hydrolytic enzymes (LytB, LytC, CbpE) (152) play an important role in pneumococcal colonization (81). Moreover, CBPs such as CbpA (Choline-Binding Protein A) and PspC are involved in adhesion (155), while PspA has the ability to bind to human lactoferrin that plays an important part in innate immunity (81).

Another surface protein that plays a role in virulence is the serine protease (PrtA), containing an LPXTG-anchor. PrtA is present in almost all the pneumococcal isolates and is an important virulence factor as strain D39 lacking PrtA was attenuated in a murine model of pneumococcal disease (28). The exact role of PrtA is not well understood, but its expression is co-regulated with other virulence genes, like the pilus genes, via by the transcriptional regulator PsaR (pneumococcal surface antigen R) in response to zinc /manganese (95, 129). PsaR is also regulating the *psaABC* genes, encoding a manganese transporter (129). The substrate-binding lipoprotein PsaA plays an important role in adhesion, as deletion of *psaA* leads to compromised adhesion and virulence, and increased oxidative stress in *S. pneumoniae* (161, 183). Other lipoproteins that play a role in virulence of *S. pneumoniae* are peptidyl-prolyl isomerases (SlrA and PpmA) (57, 97) and the lipoprotein components of iron uptake systems (PiuA and PiaA) (35). In addition, pneumococcal histidine triad (Pht) proteins (PhtA, B, D and E) were also shown to have a role in virulence as mutating any of them leads to attenuation of *S. pneumoniae* (185).

Other important virulence factors are neuraminidases (NanA and NanB) and hyaluronate lyase (HylA). NanA and NanB enhance the chances of pneumococcal colonization by degrading sialic acid present in host surface glycans, and promote the development of otitis media and formation of biofilms (26, 108, 226, 247-249). HylA degrades the hyaluronan present in the extracellular matrix of host tissues (122, 144). Thus, all these virulence factors help this human pathogen to cause morbidity and mortality by facilitating colonization of and proliferation in the host, and by providing defense against the human immune system. Several of these virulence factors are potentially good candidates for the development of protein vaccines.

Transcriptional responses of *S. pneumoniae* to a changing environment

During favorable environmental conditions, pneumococcus has the ability to spread from the nasopharynx to different parts of the human body, including the bloodstream, airways, lungs and meningeal tissues, and thereby cause pneumococcal infections. Obviously, in all these niches it encounters different environmental conditions. These environmental conditions may vary in availability of nutrients, like different carbohydrate- and nitrogen sources, micronutrients, including varying concentrations of trace metal ions, oxidative and osmotic stress, temperature and pH variation, and many others. Detailed knowledge of the impact of changing environmental conditions on *S. pneumoniae* and the expression of its

virulence factors will help not only to understand the life-style of this human pathogen, but will also help to develop smarter strategies to overcome pneumococcal infections.

Previous studies have highlighted the effects of a changing environment on the gene expression of *S. pneumoniae*, like oxidative-, osmotic- and heat-stress, acid tolerance, availability of different carbon-, amino acid- and nitrogen sources, and limitation/stress of trace metals (36, 43, 96, 105, 107, 125-127, 129, 139, 176, 223-225, 246). Inside the host, survival of pneumococcus requires the acquisition of certain nutrients like amino acids, carbon sources and metal ions. *S. pneumoniae* does have dedicated systems to fulfill its needs for amino acids, many of which are acquired from the environment (100, 126, 127). Furthermore, to sustain in its natural environment and cause infections, it requires efficient use of carbohydrates as energy sources (100, 243). *S. pneumoniae* also has dedicated systems to transport these nutrients inside the cell and some of these systems have shown to be important for pneumococcal survival inside the host.

The main objective of this thesis is to study the transcriptional responses of *S. pneumoniae* to these varying environmental conditions, especially to the availability of different carbohydrate sources and concentrations of trace metal ions, as well as the transcription factors involved.

Regulation of trace metal-ion homeostasis in *S. pneumoniae*

In recent years, several studies have shown that in pathogenic bacteria including *S. pneumoniae* there is a complex intertwinement of metabolic (regulatory) networks, the functioning and expression of virulence factors, and the general- or site-specific virulence potential (6, 7, 39, 43, 94, 95). An important class of nutritional factors that lie on the intersection between nutrient metabolism and virulence are trace metal ions (44, 123, 189, 259). In order to function well, all forms of life require trace metal ions and it is estimated that approximately 30% of all proteins and almost 50% of all enzymes require metal ions as co-factors, which determine the activity and/or the structure of these proteins (9, 237, 260). In addition, metal ions are important for the maintenance of the structure of other macromolecules like DNA and also have a role in the bacterial cell wall (260). Bacteria use various dedicated transport systems that are tightly regulated by metal-responsive transcriptional regulators to maintain an adequate intracellular level of trace metal ions (37, 55, 91, 123, 174, 189) (Fig. 3). Many of these regulators and their regulons have been demonstrated not only to be involved in metal homeostasis, but also in (the regulation of)

virulence (gene expression), as well as bacterial physiology, by affecting the functioning of metal-dependent proteins (Fig. 3).

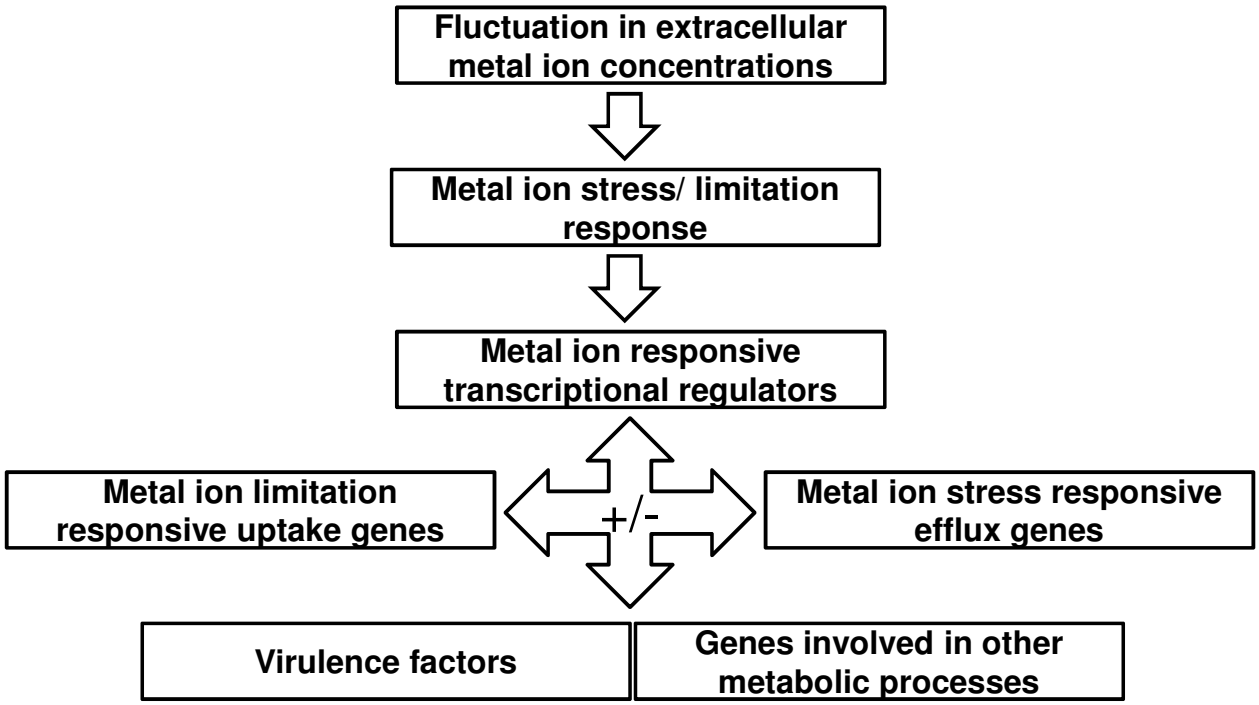


Figure 3: Flow chart representing the response of a bacterium during fluctuating metal ion concentrations.

Many years of research have focused on the importance of iron in the interplay between pathogens and their hosts (37, 90, 175, 218). Bacterial pathogens have evolved sophisticated systems to take up iron from host tissues, such as high affinity iron transport systems as well as siderophores that capture free iron with very high affinity (90, 170). In addition, bacterial cells may produce specific binding proteins to capture iron sources like heme, transferrin, lactoferrin and hemoglobin (170). The human body lowers free iron levels as much as possible to prevent bacterial growth, a process known as nutritional immunity (44, 117). Previous studies have also shown the importance of iron acquisition mechanisms for the life-style of *S. pneumoniae* (35, 37, 38). Moreover, besides iron, other metals, like copper, zinc and manganese have been shown to also contribute to the pathogenesis of streptococci including *S. pneumoniae* (12, 21, 34, 95, 110, 111, 128, 129, 166, 206, 212, 224, 225). Recently, particularly important insights have been gained into the pivotal role of zinc (12, 21, 34, 166).

Zinc is a metal ion with specific features. It is functional both as a structural component as well as a catalytic component in some proteins (207). Also, it is (like copper) a highly competitive metal ion that can replace the cognate metal ion in certain metal sensors

(85, 207). Compared to other metal ions like manganese and iron, free zinc concentrations are thought to be kept very low in the cell, due to a high cellular zinc chelating capacity (187). *S. pneumoniae* has zinc-uptake and -efflux systems that ensure optimal obtention of zinc from the surrounding environment, as well as export zinc when its concentration exceeds the tolerated levels. Zinc-uptake is of crucial importance for the virulence of *S. pneumoniae*, but on the other hand, elevated zinc levels that may be induced in the host during infection can also be detrimental for this pathogen (and others) by causing manganese restriction (11, 21, 32, 105, 129, 145, 148, 149, 166, 185, 207, 210, 224, 265).

In many bacteria, like *Escherichia coli* and *Bacillus subtilis*, Zur (Zinc uptake regulator) regulates zinc uptake genes (188). However, in *S. pneumoniae*, a different type of regulator is involved, namely the MarR family regulator AdcR (188, 224). During zinc starvation, AdcR binding to its target promoters is relieved, thereby expressing genes involved in zinc acquisition, ensuring maintenance of intracellular zinc levels (105, 224) (Fig. 4). Interestingly, AdcR regulates the genes that encode the Pht family proteins, which are believed to be surface-exposed proteins that can bind zinc (207, 224). In addition, in some streptococci, AdcR might regulate alternative ribosomal genes, which might be induced upon zinc starvation leading to lower consumption of zinc by ribosomes (12, 34, 188).

Although, *S. pneumoniae* has a dedicated regulator for zinc uptake, fulfilling a function similar as Zur in other bacteria, an interesting novel regulator is involved in zinc stress. This regulator, SczA, conserved in almost all streptococci, activates the *czcD* zinc-efflux gene upon encountering high levels of zinc, which occurs via mechanism involving both a SczA operator site that functions during low zinc levels and a SczA regulatory site that confers activation of *czcD*, needed for induction of transcription during high zinc levels (128) (Fig. 4). Besides *czcD*, SczA also induces expression of two downstream genes, *nmlR* and *adhC*. NmlR has been shown to activate *adhC* expression, which in turn has been proposed to encode a GSNO or GS-aldehyde reductase, providing resistance against nitric oxide (NO) (200, 236). Since *adhC* is predicted to contain zinc, and since reduced glutathione might be necessary for reversing zinc-binding cysteine active sites, when they are oxidized upon oxidative stress, the induction of expression of these genes at high zinc levels makes sense. Together, both zinc-dependent regulators govern the cytoplasmic levels of zinc: AdcR tunes expression of zinc acquisition systems encoded by the *adc* and *pht* genes, when zinc is lowly abundant (lower than a few μM), and SczA takes over control at higher levels of zinc by proper adjustment of expression of the CzcD zinc-efflux pump, consistent with metal-binding affinities of both zinc sensors (85).

Manganese also plays a crucial role in the virulence and gene regulation of *S. pneumoniae* (189, 212). PsaR has been shown to be responsible for manganese-dependent repression of the manganese-uptake system PsaBCA, as well as the choline binding protein PcpA and the surface associated protein PrtA. In *S. pyogenes*, *S. agalactiae* and *S. mutans*, this regulator is, besides manganese, also involved in the response to iron and seems to have an impact on other important physiological processes as well, such as biofilm formation, genetic competence, oxidative stress tolerance and adherence (33, 129). Interestingly, the regulon of PsaR is upregulated in conditions with high zinc levels (105, 129) (Fig. 4). *In vitro*, zinc ions prevent manganese-induced interaction of PsaR with its cognate promoters (129). This might resemble zinc binding to the homologous regulator MntR of *B. subtilis* that also leads to a conformational state of the protein impeding its function (79, 124). In addition, *in vivo*, zinc ions might be able to compete with manganese ions for binding to PsaA (166). This binding seems irreversible, and thereby, zinc strongly lowers intracellular manganese levels in *S. pneumoniae*, leading to a virulence phenotype similar as in a *psa* mutant (166). Recently a cation efflux system (MntE) was identified that was not only required for proper homeostasis of manganese in *S. pneumoniae*, but also for pathogenesis (107, 212) (Fig. 4).

Another important trace metal for most organisms is copper, which is toxic at high levels. Compared to blood, free levels of copper are found to be high in the lungs (lungs 121.96 and blood 12.98 $\mu g\ g^{-1}$ dry weight, (46)). Bacteria have a number of efflux systems (CopA/CopB P1-type ATPase transporters) for the removal of excessive copper and initiate a genetic response to prevent the toxicity caused by high levels of copper (18, 118, 225, 238, 261). This genetic response is controlled by copper-dependent transcriptional regulators. To date, two main types of Gram-positive copper-responsive regulators have been characterized. These are the CopY copper-responsive repressor in *Enterococcus* and *Streptococcus* spp. (197, 225), and the CsoR copper-responsive repressors found in *Mycobacterium tuberculosis* (146) and *B. subtilis* (231).

Copper homeostasis was extensively studied in *Enterococcus hirae* (206, 232, 233). In this bacterium, CopY, encoding a transcriptional regulator, is responsible for the regulation of copper-uptake/-efflux genes (cop^{EH} operon: comprised of four genes *copYZAB*) (50, 206, 233). In case of low intracellular copper, the extracellular reductase CorA supplies copper for uptake by CopA (encoding copper-transporting P-type ATPase), which further delivers it to CopZ (a copper chaperone protein) for the activation of the *cop* operon by CopY. However, in case of excessive intracellular copper, CopZ delivers copper to CopB, a copper-export ATPase. Compared to *E. hirae*, the cop^{SP} operon in *S. pneumoniae* consists of three genes,

including *copA* and *copY*, but *copZ* and *copB* are absent. The third gene, *cupA*, of the *cop*^{SP} operon encodes a putative copper binding protein. Absence of CopZB suggests that CopA may act as a copper-exporter and/or -importer. Our work showed that CopA is required not only to overcome the toxicity of copper but is also required for the virulence of *S. pneumoniae* (225).

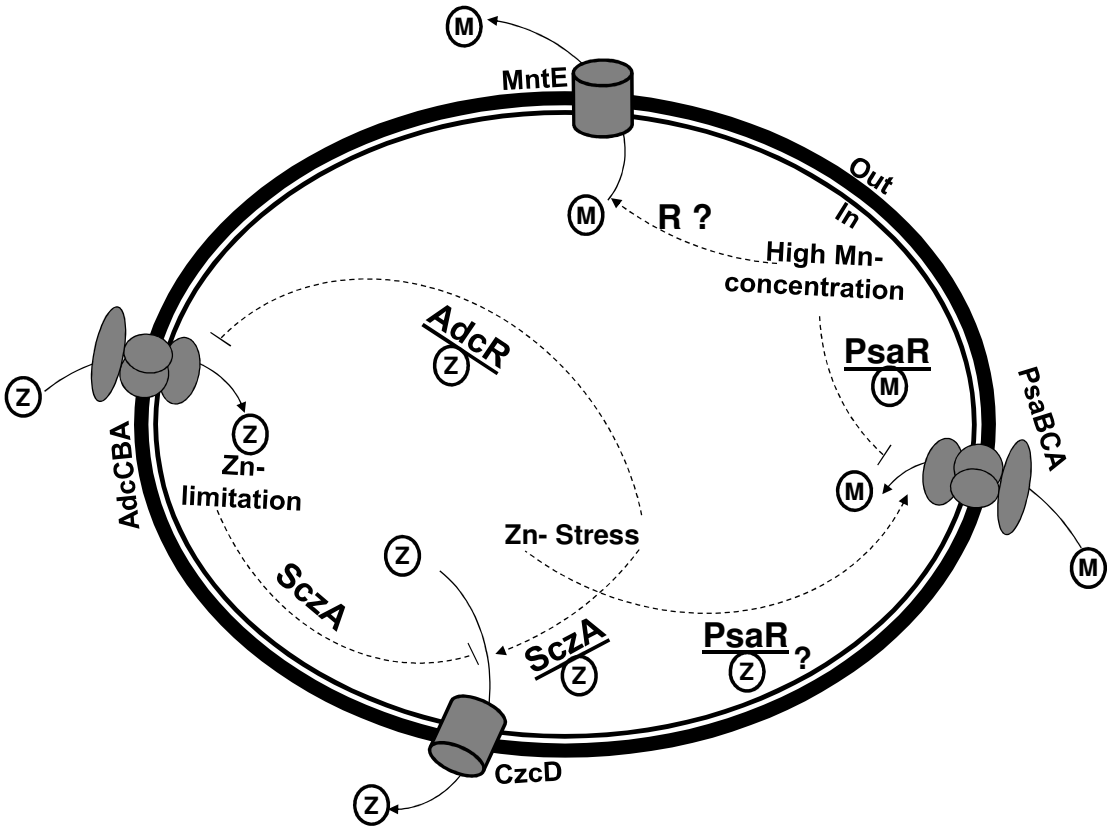


Figure 4: Graphical representation of the roles of the most important proteins in zinc (Z) and manganese (M) homeostasis in *S. pneumoniae*.

Like zinc, manganese and copper, other trace metal ions such as magnesium and iron also are likely to have an impact on gene expression in *S. pneumoniae*. Recently, a P-type ATPase transporter that has a high homology to a *Salmonella* magnesium/calcium transporter (MgtA) was characterized in *S. pneumoniae* as a magnesium/calcium transporter (176, 213). *S. pneumoniae* possesses three operons for iron transport/uptake (251). However, their regulatory mechanisms are still largely unknown. Notably, the orphan response regulator RitR and the global transcriptional regulator CodY were shown to regulate one of the iron uptake systems independently of the iron concentration (47, 251, 252). Therefore, there are many unknown horizons that still need to be explored in the area of metal-dependent gene regulation in *S. pneumoniae*.

Regulation of carbohydrate metabolism in *S. pneumoniae*

The survival of the human pathogen *S. pneumoniae* is not only dependent on proper regulation of virulence factors, but also on the adaptation and utilization of available carbohydrates inside the human body. There are various kinds of carbohydrate (derivatives) that it may encounter in its natural environment. However, the exact sources of carbohydrates exploited by pneumococcus in natural environment are not known. More than 30% of the transporters in *S. pneumoniae* are predicted to be sugar transporters (29, 40, 243). These transporters include classical PTSs (phosphoenolpyruvate (PEP)-dependent phosphotransferase systems), ATP-binding cassette (ABC) and ion-gradient-driven transporters (29, 40). The number of PTSs varies from 15 to 20 in different pneumococcal strains (29, 40). These PTSs are basically comprised of two cytoplasmic components EI (Enzyme I) and HPr that are common in all PTSs, and EII (Enzyme II), which has carbohydrate specific residues (68, 199).

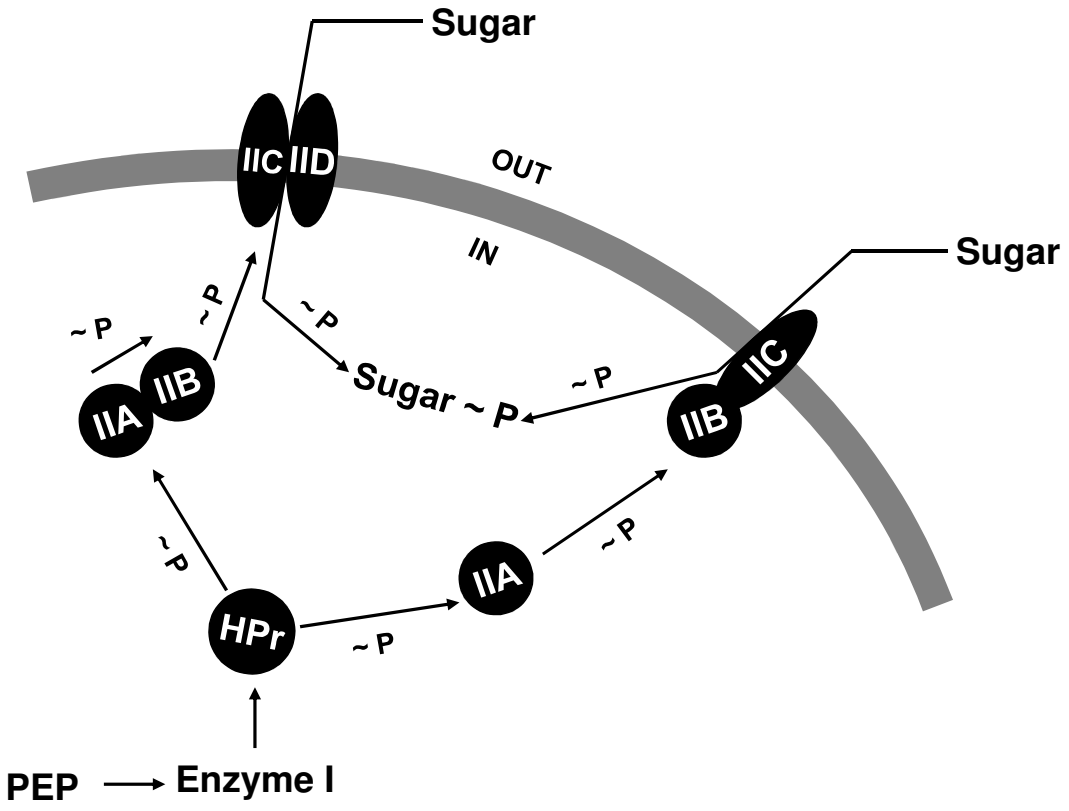


Figure 5: Graphical representation of bacterial phosphoenolpyruvate-dependent phosphotransferase systems.

The basic composition of PTSs is almost the same in all species studied so far (20, 68, 199, 211). Bacteria usually have many EIIs that are specific to various carbohydrate sources. The EII complexes consist of one or two hydrophobic integral membrane domains (C and D)

and two hydrophilic domains (A and B) (68). These domains can be encoded as a single protein or up to four separate proteins (217). These domains together are involved in the carbohydrate transport as well as in phosphorylation of the carbohydrate upon entering the cell (68). This phosphorylation takes place by a multiple protein phosphoryl-transfer chain, which is initiated by donating phosphate to EI by PEP (Fig. 5). EI then phosphorylates HPr, which transfers this phosphate to the EII component A/B of the PTS. Finally, EIIA/B phosphorylate the incoming sugar to sugar-6-P, whereas EIIC/D are responsible for the entry of the sugar inside the cell. *S. pneumoniae* strain D39 has around 20 PTS sugar-specific EII complexes with different domain arrangements (103). Table 1 summarizes all the putative PTSs present in *S. pneumoniae* D39 as well as the putative transcriptional regulators that could be involved in regulation of these PTSs.

Like for many other bacteria, glucose is the preferred source of energy for *S. pneumoniae* (29, 103). Compared to human blood (3.57-6.06 mM), glucose levels at common sites of pneumococcal infection are generally quite low (saliva 0.02-0.4 mM, nasal secretion <1.0 mM, lower airway secretions <0.5mM) (60, 82, 192, 228, 270), suggesting that pneumococcus should have a good ability to utilize other carbohydrate derivatives available in these places (119, 121). The ability to use a preferred sugar before the utilization of a non-preferred sugar depends on a regulatory process called carbon catabolite repression (CCR) (23, 43, 80, 101, 103). CCR is involved in the silencing of genes that are specific to utilize non-preferred sugars, until the cell has consumed the preferred sugar(s) (23, 43, 77, 103). This CCR is controlled by carbon catabolite control protein A (CcpA) in many low GC Gram-positive bacteria including *S. pneumoniae* and it helps the bacteria to efficiently adapt according to the available carbon source(s) (68, 80). CcpA belongs to the LacI/GalR family of transcriptional activators/repressors and binds to catabolite responsive element (*cre*) located within or near promoter regions of genes that are under the control of CcpA (43, 154, 281). The position of the *cre* box determines the role of CcpA as an activator or a repressor of targeted genes (43, 154, 281). If a *cre* box is located within a promoter region or open reading frame (ORF), binding of CcpA inhibits RNA polymerase interaction with the promoter or its processivity over the DNA, thereby repressing the transcription of the gene (43, 154, 281). On the other hand, binding of CcpA to a *cre* box located upstream of the promoter enhances the transcription by allowing CcpA to stabilize the RNA polymerase-promoter interaction (43, 154, 281). Moreover, CcpA binding to a *cre* box is enhanced by the binding of Ser46-phosphorylated HPr, a component of the multi-protein PEP-dependent PTS, to CcpA (68).

The role of the proper utilization of different sources of carbohydrates in the fitness and virulence of *S. pneumoniae* has been investigated previously (29, 43, 48, 77, 103, 104, 144, 159, 160, 168, 179, 214, 223, 241). These carbohydrate sources include maltose, mannose, galactose, cellobiose, raffinose, sialic acid, sucrose and others. Expression of several of these carbohydrate metabolic genes is controlled by specific transcriptional regulators, either or not in addition to CCR. These include regulators from the LacI/GalR, GntR, ROK and other families. Malto-oligosaccharide utilization is repressed by the LacI/GalR family repressor, MalR (29, 179, 180, 202). MalR directly binds to promoters of two operons that are involved in transport and metabolism of malto-oligosaccharides, and represses their expression in the absence of malto-oligosaccharides (179, 180, 202). *S. pneumoniae* also has the ability to transport sucrose by ABC transporters, the corresponding genes of which are repressed by the LacI/GalR family repressor ScrR (104). Another LacI/GalR family regulator studied in *S. pneumoniae* is RegR. Initially, RegR was predicted to be a global transcriptional regulator of carbohydrate utilization in *S. pneumoniae*, but is actually a specific regulator for hyaluronic acid degradation and utilization (49, 255).

Recently, the raffinose and cellobiose utilization loci have also been characterized in *S. pneumoniae* (168, 214, 223, 250). The raffinose utilization locus was first characterized in R6 (214) and recently in the D39 strain (250). The expression of this locus is controlled by both a transcriptional activator RafR and a transcriptional repressor RafS (214, 250). On the other hand, CelR, a cellobiose-dependent transcriptional activator has been described to mediate the expression of a putative cellobiose utilization gene cluster (*cel* locus) in *S. pneumoniae* (223). The recently published RegPrecise regulon database indicates the presence of 11 different carbohydrate regulators including the above mentioned ones (181). These regulators are predicted to control the expression of genes involved in the metabolism of fucose, fructose, lactose, galactose, NAG, trehalose and other sugars. A comprehensive study, showing the role of these predicted regulators in the regulation of different sugar uptake systems should be performed in *S. pneumoniae*. This will give more insight into how *S. pneumoniae* adapts to varying carbohydrate availability.

Table 2: List of PTS systems in *S. pneumoniae* D39 strain (29, 40)**.**

S/N	D39 locus tag	Domain	PTS family	Predicted substrate	Putative Regulators	References
	SPD1039/40	**EI/HPr**		**General PTS**	**CcpA**	(43)
1	**SPD0066/67/68/69**	**IIB/IIC/IID/IIA**	**Man**	**Mannose, Galactose**	**SPD0064**	(109, 241)
2	**SPD0232/33/34**	**IIA/IIB/IIC**	**Lac**	**-**	**-**	(29, 40)
3	**SPD0262/63/64**	**IID/IIC/IIAB**	**Man**	**Glucose, Mannose, Fractose, GlcNAc**	**-**	(45, 52, 153)
4	**SPD0279/81/83**	**IIB/IIA/IIC**	**Lac**	**Beta-glucosides**	**SPD0280 (CelR)**	(Chapter 4)
5	**SPD0293/95/96/97**	**IIA/IIB/IIC/IID**	**Man**	**Hyaluronic acid**	**SPD0301 (RegR)**	(144, 163)
6	**SPD0360/62**	**IIC/IIA**	**Gat**	**Manitol**	**SPD0361 (MtlR)**	(99)
7	**SPD0424/26/28**	**IIC/IIA/IICB**	**Lac**	**Mannose, Lactose, Cellobiose**	**SPD0423 (RokA)**	(Chapter 6)
8	**SPD0502**	**IIBCA**	**Glu**	**Beta-glucosides**	**SPD0500**	(56)
9	**SPD0559/60/61**	**IIA/IIB/IIC**	**Gat**	**Lactose, Galactose, Galactitiol**	**SPD1049 (LacR)**	(115)
10	**SPD0661**	**IICBA**	**Glc**	**Maltose**	**SPD1938 (MalR)**	(138, 227, 263)
11	**SPD0773**	**IIABC**	**Fru**	**Fructose**	**SPD0771**	(131)
12	**SPD1047/48**	**IIA/IICB**	**Lac**	**Lactose, Tagatose**	**SPD1049 (LacR)**	(1, 63, 277, 278)
13	**SPD1057**	**IIB**	**Gat**	**-**	**-**	
14	**SPD1496**	**IICB**	**Glc**	**Aminosugars**	**SPD1488**	(29)
15	**SPD1532**	**IIBCA**	**Glc**	**Sucrose**	**SPD1535 (ScrR)**	(104)
16	**SPD1664**	**IIBCA**	**Glc**	**Trehalose**	**SPD1665 (TreR)**	(29)
17	**SPD1831/32/33**	**IIC/IIA/IIB**	**Lac**	**Beta-glucosides**	**SPD1829 (CelR-II)**	(Chapter 5)
18	**SPD1845/46/47**	**IIA/IIB/IIC**	**Asc**	**Ascorbate**	**SPD1841**	(102, 280)
19	**SPD1959/60**	**IIC/IIB**	**Asc**	**Xylulose, Ribulose**	**SPD1961**	(29)
20	**SPD1989/90/91/92**	**IID/IIC/IIB/IIA**	**Man**	**Arabinose, Fucose**	**SPD1996 (FucR)**	(48, 98)

Outline of the thesis

Changing environmental factors like varying metal ion concentrations and carbohydrate availability may have a huge impact on the life style and the virulence of *S. pneumoniae*. This thesis provides a increased knowledge on the regulatory responses of this human pathogen to some environmental changes in nutrient levels that may occur inside the human body.

Chapters 2 and **3** elaborate on the effect of zinc and copper concentrations on gene expression in *S. pneumoniae*. **Chapter 2** describes the transcriptional response of *S. pneumoniae* to zinc-limiting conditions that result in upregulation of the AdcR regulon, and shows that AdcR functions both as a transcriptional repressor and activator. The AdcR regulon includes important virulence factors belonging to the Pht family proteins and the zinc uptake *adc* operon. **Chapter 3** describes the transcriptomic response of *S. pneumoniae* to copper. It is shown that high copper levels result in increased expression of the *cop* operon, which is regulated by the transcriptional regulator CopY. The expression of the *cop* operon was also affected by increasing concentrations of zinc. Additionally, the *cop* operon was induced in the lungs and nasopharynx of intranasally infected mice and a *copA*$^-$ mutant strain that has decreased growth at high levels of copper *in vitro* showed reduced virulence in a mouse model of pneumococcal pneumonia.

The effect of cellobiose on gene regulation of *S. pneumoniae* is elaborated in **Chapters 4** and **5**. **Chapter 4** highlights the functional characterization of the cellobiose-dependent transcriptional activator CelR. Transcriptomic comparison of a *celR* mutant and D39 wild-type allowed identification of specific CelR targets (*i.e.* cellobiose utilization gene cluster *cel*). Additionally, a regulatory site of CelR was identified by promoter truncation experiments. This regulatory site is highly conserved in other streptococci as well, which suggests a similar role of CelR in other streptococci. **Chapter 5** describes the genome wide transcriptional response of *S. pneumoniae* to cellobiose. Expression of many genes with diverse functions was affected in the presence of cellobiose, including the *cel* gene cluster (discussed in **Chapter 4**) and the *cel-II* operon. Further functional characterization of *celR-II*, a GntR transcriptional regulator that acts as a transcriptional repressor for the *cel-II* operon, was performed in **Chapter 5**.

The role of a ROK family transcriptional repressor, RokA, is described in **Chapter 6**. *S. pneumoniae* has four proteins belonging to the ROK family and only one of them (RokA) has a HTH motif. A transcriptomic comparison of a *rokA* mutant and wild-type strains has

revealed high expression of an operon (*SPD0424-28*) that is putatively involved in carbohydrate metabolism. Moreover, a DNA operator site for RokA was identified by promoter truncation experiments, which was further confirmed by electrophoretic mobility shift assays and DNA footprints using a purified strep-tagged derivative of RokA.

Chapter 7 summarizes the results and describes the prospects of the research performed in this thesis.

In conclusion, this thesis contributes to a further understanding of the role of metal ion homeostasis- and carbohydrate metabolism, and the regulation of these processes, in the human pathogen *S. pneumoniae* and its virulence.

Chapter 2

Transcriptional response of *Streptococcus pneumoniae* to Zn^{2+} limitation and the repressor/ activator function of AdcR

Sulman Shafeeq, Tomas G. Kloosterman and Oscar P. Kuipers

Based on *Metallomics* **3(6)**: 609-18 (2011).

Abstract

Zinc (Zn^{2+}) is an important trace metal ion that has been shown to regulate the expression of several (virulence) genes in streptococci. Previously, we analyzed the genome-wide response of *S. pneumoniae* to Zn^{2+}-stress. In this work, we have performed a transcriptomic analysis to identify genes that are differentially expressed under intracellular Zn^{2+} limitation. This revealed a number of genes that are highly upregulated in the absence of extracellular Zn^{2+}, amongst which the genes belonging to the regulon of the Zn^{2+}-responsive repressor AdcR, like *adcBCA*, encoding a Zn^{2+}-dependent ABC-uptake system, *adcAII,* encoding a Zn^{2+}-binding lipoprotein, and also virulence genes belonging to the Pht family (*phtA*, *phtB*, *phtD* and *phtE*). Using transcriptome analysis, *lacZ*-reporter studies, *in vitro* DNA binding experiments, and *in silico* operator predictions, we show that AdcR directly represses the promoters of *adcRCBA*, *adcAII-phtD*, *phtA*, *phtB* and *phtE* in the presence of Zn^{2+}. AdcR can also function as an activator, since in the presence of Zn^{2+} it directly induces expression of *adh* that encodes a Zn^{2+}-containing alcohol dehydrogenase. In conclusion, the genome-wide transcriptional response of *S. pneumoniae* to Zn^{2+} limitation was established, which is mainly mediated via direct regulation by the Zn^{2+}-dependent regulator AdcR.

Introduction

Streptococcus pneumoniae is an important Gram-positive human pathogen that causes millions of deaths each year worldwide (167). It usually resides asymptomatically in the nasopharynx (31), but occasionally it may spread to the lungs, blood, meninges and sinuses, causing serious infections like pneumonia, otitis media, septicaemia or meningitis (114, 182). During infection, *S. pneumoniae* encounters many environmental changes, which may affect the expression of virulence and other genes. In previous studies, it has been shown that trace metal ions can be important factors in the virulence and physiology of several human pathogens (8, 30, 35-37).

Trace metal ions are essential for the proper functioning of many enzymes and transcriptional regulators in bacteria. However, high concentrations of trace metal ions can be lethal to the cell (30, 73, 174, 178). Therefore, homeostasis of trace metal ions is vital for the survival of bacterial pathogens (30, 73). *S. pneumoniae* possesses many trace metal ion acquisition- and efflux systems (35, 36, 70, 128, 129). These systems ensure the maintenance of proper levels of trace metal ions in the cell and their expression may be tightly controlled by metal-responsive regulatory proteins (128, 129).

Zinc (Zn^{2+}) is likely to be an important environmental factor for *S. pneumoniae* as its concentration varies from a few µM to over 100 µM in the human body (258), it has been shown to affect the expression of several genes (129), and it is also a co-factor of many proteins (53). As Zn^{2+} is toxic at high concentrations, it is important to maintain a proper Zn^{2+} homeostasis in the cell. So far, three Zn^{2+}-dependent regulatory proteins have been characterized in *S. pneumoniae*, namely PsaR, SczA and AdcR (69, 70, 105, 128, 129, 185, 207). PsaR belongs to the DtxR family of regulators and regulates the genes encoding the PsaBCA Mn^{2+} transporter and the choline binding protein PcpA in a Mn^{2+}- and Zn^{2+}-dependent way (129). SczA is a TetR family regulator, more specifically a $Zn^{2+}/Co^{2+}/Ni^{2+}$-dependent transcriptional activator of the Zn^{2+}-resistance gene *czcD* necessary for *S. pneumoniae* to cope with Zn^{2+}-stress conditions (128).

AdcR is a MarR family regulatory protein and a repressor of the *adcRCBA* operon, encoding a Zn^{2+}- and/or Mn^{2+}- dependent ABC transporters, the genes encoding Pht family proteins and *adcAII* (10, 185, 207). These genes share the presence of a possible AdcR DNA binding site in their promoter regions (10, 188). Recently, it has been shown in *S. pneumoniae* that AdcR is a homodimer in solution, and several histidine residues have been shown to be essential for the Zn^{2+}-dependent binding of AdcR to the *adcR* promoter region (207).

The *adc* operon has been studied in other streptococci as well. In *Streptococcus gordonii*, the *adc* operon is found to be responsible for Mn^{2+} homeostasis (151) and mutation of *adcR* delays biofilm formation (173). In *Streptococcus suis*, AdcR is involved in the regulation of genes encoding a Zn^{2+}-/Mn^{2+}-uptake system, ribosomal genes and *pht* genes, and it was shown to require Zn^{2+}/Mn^{2+} to bind with some of its target promoters (10). Moreover, it has been found that AdcR is required for full virulence and protection against oxidative stress in *S. suis* (11).

In a previous study, we have determined the effect of Zn^{2+}-stress on the transcriptome of *S. pneumoniae* (129). In this study, we investigate the transcriptional response of *S. pneumoniae* to Zn^{2+}-limitation using DNA microarrays. We show that AdcR is directly responsible for most of the Zn^{2+}-dependent effects that are observed in the DNA microarray studies. Furthermore, AdcR is found to function both as an activator and as a repressor.

Materials and Methods

Bacterial strains, growth conditions and DNA manipulation

Table 1: List of strains and plasmids used in this study.

Strain/plasmid	Description	Source
S. pneumoniae		
D39	Serotype 2 strain, *cps 2*	Laboratory of P. Hermans
D39*nisRK*	D39 Δ*bgaA*::*nisRK*; $Trmp^R$	(125)
SS200	D39 Δ*adcR*; Ery^R	This study
SS201	D39 Δ*bgaA*::P*adcR-lacZ*; Tet^R	This study
SS202	D39 Δ*bgaA*::P*adcAII-lacZ*; Tet^R	This study
SS203	D39 Δ*bgaA*::P*PhtA-lacZ*; Tet^R	This study
SS204	D39 Δ*bgaA*::P*phtB-lacZ*; Tet^R	This study
SS205	D39 Δ*bgaA*::P*phtE-lacZ*; Tet^R	This study
SS206	D39 Δ*bgaA*::P*adh-lacZ*; Tet^R	This study
SS207	SS200 Δ*bgaA*::P*adcR-lacZ*; Tet^R	This study
SS208	SS200 Δ*bgaA*::P*adcAII-lacZ*; Tet^R	This study
SS209	SS200 Δ*bgaA*::P*phtA-lacZ*; Tet^R	This study
SS210	SS200 Δ*bgaA*::P*phtB-lacZ*; Tet^R	This study
SS211	SS200 Δ*bgaA*::P*phtE-lacZ*; Tet^R	This study
SS212	SS200 Δ*bgaA*::P*adh-lacZ*; Tet^R	This study
L. lactis		
NZ9000	MG1363 Δ*pepN::nisRK*	(134)
E. coli		
EC1000	Km^R; MC1000 derivative carrying a single copy of the pWV1 *repA* gene in *glgB*	Laboratory collection
Plasmids		
pPP2	Amp^R Tet^R; promoter-less *lacZ*. For replacement of *bgaA* with promoter *lacZ*-fusion. Derivative of pTP1	(88)
pNZ8048	Cm^R: Nisin inducible P*nisA*	(62)
pSS101	pPP2 P*adcR-lacZ*	This study
pSS102	pPP2 P*adcAII-lacZ*	This study
pSS103	pPP2 P*phtA-lacZ*	This study
pSS104	pPP2 P*phtE-lacZ*	This study
pSS105	pPP2 P*phtB-lacZ*	This study
pSS106	pPP2 P*adh-lacZ*	This study
pSS107	pNZ8048 carrying strep-tagged AdcR in downstream of P*nisA*	This study

Bacterial strains and plasmids constructed in this study are listed in Table 1. Strains were stored in 10% glycerol at -80° C. *S. pneumoniae* was grown as described (125). Where indicated, chemically defined medium (CDM) with a pH of 6.4 was prepared as described (125, 126), with the exception that $ZnSO_4$ was omitted from the metal mixture and added separately as specified in the results section. Metal ions were added as the salts of $ZnSO_4$, $MnSO_4$, $MgCl_2$, $CaCl_2$, $CoCl_2$, $NiSO_4$, $CuSO_4$, $FeCl_3$ and $FeCl_2$. *Escherichia coli* strain

EC1000 and *Lactococcus lactis* strain NZ9000 were cultured at 37 and 30 °C, respectively. Where necessary for selection, media were supplemented with the following concentrations of antibiotics: erythromycin: 0.25 μg ml^{-1} for *S. pneumoniae* and 120 μg ml^{-1} for *E. coli*; spectinomycin: 150 μg ml^{-1} for *S. pneumoniae*; tetracyclin: 2.5 μg ml^{-1} for *S. pneumoniae*; chloramphenicol: 4 μg ml^{-1} for *L. lactis*; and ampicillin: 100 μg ml^{-1} for *E. coli*. When appropriate, nisin (sigma) was used to induce nisin-dependent overexpression at a concentration of 10 ng ml^{-1} in *L. lactis* (126). Primers used in this study are listed in Table 2. Chromosomal DNA of *S. pneumoniae* wild-type strain D39 was used as a template for PCR amplification (15, 136) and all DNA manipulations were done as described (125).

Table 2: List of primers used in this study.

Name	Nucleotide sequence (5' to 3')[a]	Restriction site
Padcr-F	CGGAATTCTTTTTCAGCAAAGATTGGG	EcoRI
Padcr-R	CGGGATCCCTTTCCTTTTAGACTTCTC	BamHI
PadcAII-F	CGGAATTCCTTCACTTATGGCTATAAGC	EcoRI
PadcAII-R	CGGGATCCAAAGAAAGACACTTAACAGG	BamHI
PphtA-F	CGGAATTCTGAACTTCAAAAAGAATACG	EcoRI
PphtA-R	CGGGATCCCTTAAAATCAAAGCTGCCGC	BamHI
PphtB-F	GCATGAATTCGGCAGAAGCAGAAAAATTAC	EcoRI
PphtB-R	CGATGGATCCAAGTGTAGCTACTGACC	BamHI
PphtE-F	CGGAATTCAGAAGTAGATAGTCTCTTGG	EcoRI
PphtE-R	CGGGATCCACGATAACAGCTGATCCAGC	BamHI
Padh-F	CATGGAATTCACATCGATGACGTCTGTGG	EcoRI
Padh-R	CATGGGATCCACAAATAGTGGTTTTTACAATAC	BamHI
PadcR -1	GAGCAGATCAAGAAGGATTTGAAG	-
PadcR -2	GCATGGCGCGCCTAGCTGTCTCATATCTGACTCC	AscI
PadcR -3	GCATGCGGCCGCGTAGGAGAAATCAAATAATG	NotI
PadcR -4	CAACGCCATGGCGAGTCTTGG	-
Ery-F	GCATGGCGCGCCATGAACGAGAAAAATATAAAACACAGTC	AscI
Ery-R	GCATGCGGCCGCCAATAGCTATAAATTATTTAATAAGTAA	NotI
PadcR-nco	CATGCCATGGCTTAAAAAAGAGAAGTCTAAAAG	NcoI
PadcR- Cstrep-xba	TGCTCTAGATTATTTTTCAAATTGTGGATGGCTCCAAGCGC TTTTGATTTCTCCTACTAAAGC	XbaI

[a]Restriction enzyme sites are underlined.

Construction of the *adcR* mutant

The *adcR* (*SPD2000*) mutant was made by allelic-replacement with an erythromycin-resistance cassette. In short, primers adcR-1/adcR2 and adcR-3/adcR-4 were used to generate PCR products of the approximately 600 bp left and right flanking regions of *adcR*, which were, by means of ligation using AscI/NotI restriction sites, fused to an erythromycin-resistance gene PCR amplified with primers Ery-F and Ery-R from plasmid pORI28 (140). The resulting ligation mixture was transformed to *S. pneumoniae* D39, yielding strain SS200, and the mutation was verified by PCR and sequencing.

Inductively coupled plasma- mass spectrometry (ICP-MS) analysis

To determine the intracellular concentration of Zn^{2+} in *S. pneumoniae* strain D39, cultures were grown in 100 ml of CDM with and without 0.2 mM Zn^{2+} till an OD600 of 0.3. Cultures were centrifuged and washed (at 4 °C) once with the CDM medium and two times with phosphate-buffered saline (PBS) that had been treated with chelex (Sigma) for overnight. The cell pellets were dried overnight in a Speedvac. The dried cells were subsequently used for analysis by means of ICP-MS, as described before (105). Results were expressed as μg of Zn^{2+} g^{-1} dry weight of cells.

Construction of *lacZ*-fusions and β-galactosidase assays

Chromosomal transcriptional *lacZ*-fusions to the *adcR*, *adcAII*, *phtA*, *phtB*, *phtE*, and *adh* promoters were constructed in the integration plasmid pPP2 with the primer pairs listed in Table 2, resulting in plasmids pSS101-06. These plasmids were introduced into wild-type D39 and SS200, resulting in strains SS201-12. All plasmid constructs were checked by sequencing, and new loci created with these plasmids were verified by PCR. For β-galactosidase assays cells were grown in triplicate in CDM with and without 0.2 mM Zn^{2+} at 37 °C and harvested in the mid-exponential phase of growth. Specific β-galactosidase activity was measured as described (93).

Overexpression and purification of AdcR

To overexpress C-terminally strep-tagged AdcR, the *adcR* gene was PCR amplified with primers adcR-nco/adcR-cstrep-xba. This PCR product was restricted with NcoI/XbaI and cloned into the NcoI/XbaI sites of pNZ8048 resulting in plasmid pSS107. Overexpression of C-terminally strep-tagged AdcR was achieved in strain *L. lactis* NZ9000 (134). Cells were induced with 10 ng ml^{-1} nisin in a 1 L culture at an OD600 of 0.5 and harvested after two hour of induction. Purification of AdcR was performed using the Streptactin column from IBA according to the supplier's instructions (www.iba-go.com). Buffers without EDTA were used and the purified protein was stored at a concentration of 0.2 mg ml^{-1} in the elution buffer (100 mM Tris-HCl pH 8, 150 mM NaCl, 2.5 mM desthiobiotin, 1 mM β-mercapthoethanol) with 10% glycerol at -80 °C.

Electrophoretic mobility shift assays

Electrophoretic mobility shift assays (EMSAs) were performed as described (129). In short, PCR products of P*adcR*, P*adcAII*, P*phtA*, P*phtB*, P*phtE*, P*adh*, P*SPD1829* were labeled

with [γ-^{33}P] ATP probes. EMSAs were performed in buffer containing 20 mM Tris-HCl (pH 8.0), 5 mM $MgCl_2$, 0.1 mM dithiotreitol (DTT), 8.7% (w/v) glycerol, 62.5 mM KCl, 25 μg ml^{-1} bovine serum albumin, 25 μg ml^{-1} poly(dI-dC) and 5,000 cpm of [γ-^{33}P] ATP-labeled PCR product. Various metal ions and EDTA were added in concentrations as specified in the Results. Reactions were incubated at 30 °C for 30 min. before loading on 6 % poly-acrylamide gels. Gels were run in 0.44 M Tris-borate buffer (pH 8.3) at 100V for 90 minutes.

Transcriptome analysis using *S. pneumoniae* whole genome DNA microarrays

For DNA microarray analysis of the response to Zn^{2+}-limitation, the transcriptome of *S. pneumoniae* wild-type strain, grown in 3 biological replicates in CDM without $ZnSO_4$ was compared to the transcriptome of the same strain grown in 3 biological replicates in CDM + 0.2 mM $ZnSO_4$. Cells were harvested at an OD600 of 0.3 by means of centrifugation for 1 min at 10,000 rpm at room temperature and immediately frozen in liquid nitrogen. Experiments were performed essentially as described (126, 128).

For transcriptome analysis of *S. pneumoniae* wild-type and its isogenic *adcR* mutant (SS200), cells were grown in 3 biological replicates in CDM with 0.2 mM $ZnSO_4$ and harvested at an OD600 of approximately 0.3. All other procedures regarding microarrays were done as described (126).

DNA microarray data analysis

DNA microarray data were analyzed as described (126). In short, slide images were analyzed using ArrayPro 4.5 (Media Cybernetics Inc., Silver Spring, MD) and further processing and normalization (LOWESS spotpin-based) of slides was done with the in-house developed *MicroPrep* software. DNA microarray data were obtained from three independent biological replicates hybridized to three glass slides, of which one was a dye-swap. Expression ratios of mutant strain over the wild-type strain were calculated from the measurements of at least 5 spots. Differential expression tests were performed on expression ratios with a local copy of the Cyber-T implementation of a variant of the *t*-test. False discovery rates (FDRs) were calculated as described (253). A gene was considered differentially expressed when $p < 0.001$ and FDR < 0.05 and when at least 5 measurements were available. The DNA microarray data are available at GEO (GSE29236).

Results

Identification of Zn^{2+}-regulated genes in *S. pneumoniae*

To obtain insight in the response of *S. pneumoniae* to a limiting Zn^{2+} concentration, the wild-type strain was grown in chemically defined medium (CDM) (125) either without Zn^{2+} or with a 0.2 mM concentration of $ZnSO_4$. By means of atomic absorption spectroscopy no Zn^{2+} could be detected in the medium without added Zn^{2+}, meaning that the Zn^{2+} concentration was below the lower limit of detection (0.1 μM), which therefore represented a severe Zn^{2+}-limited condition. Table 3 summarizes the transcriptome changes induced by Zn^{2+} limitation compared to the situation with 0.2 mM Zn^{2+}. 49 genes were differentially expressed, of which 26 were up- and 23 were downregulated under Zn^{2+} limitation. The *adcRCBA* operon was highly upregulated in the absence of Zn^{2+}. This operon encodes a Zn^{2+}- and/or Mn^{2+}-dependent ABC transporters and the transcriptional regulator AdcR (69, 70). Furthermore, several genes encoding members of the family of pneumococcal histidine triad proteins (Pht) were highly upregulated. These proteins have been suggested to have a role in Zn^{2+}-acquisition as well (209, 210). Another interesting gene is *SPD0888*, encoding AdcAII, a putative laminin binding protein belonging to the LraI-lipoprotein family, which has been shown to be involved in adhesion and invasion of host cells in *S. agalactiae* and *S. pyogenes* (72, 235, 239, 240). Furthermore, members of the LraI-family are often involved in divalent cation transport (106) and recent structural studies strongly suggest that *adcAII* has a role as a cell surface Zn^{2+}-binding protein in *S. pneumoniae* (149). The *adcRCBA*, *adcAII-phtD*, *phtA*, *phtB* and *phtE* operons/genes were recently shown to be upregulated in a mutant of *adcR* (207). It is therefore likely that AdcR mediates part of the response of *S. pneumoniae* to Zn^{2+} limitation, by releasing repression in the absence of Zn^{2+}.

Two other transcriptional regulator genes, *SPD0633* (*copY*) and *SPD1594*, were also upregulated in the absence of Zn^{2+}. *CopY* encodes a copper-responsive regulator and lies in an operon with *SPD0634* and *SPD0635*, encoding proteins with homology to cupredoxin/copper-transporting P-type ATPase CupA and to the copper-influx protein CopA, respectively (225). The second transcriptional regulator gene, *SPD1594*, is in an operon with *SPD1595*, encoding a hypothetical protein. Interestingly, this regulator has homology with a Zn^{2+}-binding Cro/CI family transcriptional regulator in *Enterococcus faecalis* and *Lactobacillus* species.

The *adh* (*SPD1865*) gene, which encodes a Zn^{2+}-containing alcohol dehydrogenase, was significantly downregulated. Adh proteins are usually known to bind with Zn^{2+} (32). As expected, *czcD* was downregulated, as was the *nrdD* operon, which is in agreement with our

previous studies (32, 128). Several hypothetical proteins were also among the differentially expressed genes. Furthermore, an operon encoding fuculose metabolism was upregulated. This operon is located downstream of the *adc* operon.

S. pneumoniae D39 accumulates less Zn^{2+} during Zn^{2+}-limited growth

To determine whether the transcriptional effects observed above correlate with a lowered cell-associated concentration of Zn^{2+} ICP-MS analysis was performed on cells grown in CDM medium without and with 0.2 mM added Zn^{2+}. This analysis revealed that *S. pneumoniae* D39 grown in the absence of added Zn^{2+} has a 5-6 fold lower cell-associated amount of Zn as compared to D39 grown in 0.2 mM Zn^{2+} (12 μg g-1 dry mass of cells *versus* 68 μg g-1 dry mass of cells). The levels of Cu, Fe, Co and Ni were all just above or just below the detection limit (approximately 5 ng g-1 dry weight). No change in the cell-associated concentration of Mn was observed (45 μg g-1 dry weight). These results suggest that the transcriptomic effects as described above are mainly induced by a difference in the intracellular concentration of Zn^{2+} in these conditions. On the basis of previous studies (12, 105, 128, 129, 185, 207), we next investigated the role of the Zn^{2+}-sensing regulator AdcR in relation to the transcriptional effects seen under Zn^{2+} limitation.

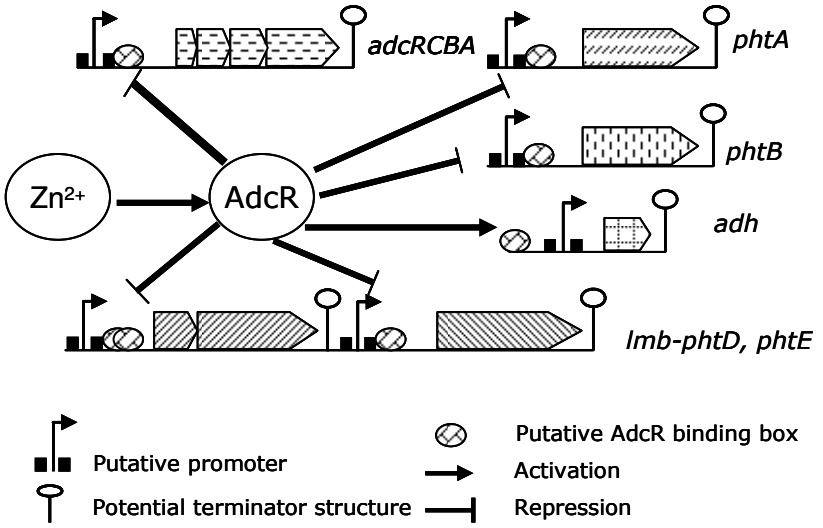
Prediction of AdcR binding sites in the promoters of the Zn^{2+}-dependent genes

The organization of the Zn^{2+}-dependent genes found in the Zn^{2+} limitation microarray is drawn in Fig. 1. In the upstream regions of *adcR*, *phtA*, *phtB, phtD*, *phtE* and *adcAII*, putative binding sites (5'-TTAACYRGTTAA-3') for the metal-responsive repressor AdcR can be found (Fig. 1) (12, 188). To find out whether *adh*, which was downregulated in the microarray, also has a putative AdcR binding site, its promoter region was analyzed by means of Genome 2D (16). A putative AdcR binding site (5'-TTTACTGGTAAA-3') was found in the *adh* promoter, suggesting that it is directly regulated by AdcR. Furthermore, our bioinformatics analysis indicates that the position of the AdcR binding site determines the role of AdcR as activator/repressor (Fig. 1A), since in all upregulated genes, a putative AdcR operator was found downstream of or overlapping with the core promoter, whereas in the downregulated gene *adh* the putative AdcR binding site is located upstream of the core promoter.

Table 3: Summary of transcriptome comparison of *S. pneumoniae* wild-type grown in CDM plus 0 mM Zn^{2+} and CDM plus 0.2 mM Zn^{2+}. [a]Gene numbers refer to D39 locus tags. [b]D39 annotation/TIGR4 annotation (100, 136, 243), [c]Ratios >2.0 or <-2.0 (wild-type + 0 mM Zn^{2+}compared to wild-type + 0.2 mM Zn^{2+}). In some cases neighbouring genes with lower ratios are also indicated.

Gene tag[a]	Function[b]	Ratio[c]
SPD0187	Anaerobic ribonucleoside-triphosphate reductase NrdD	-2.2
SPD0188	Predicted acetyltransferase, GNAT family	-2.2
SPD0189	Anaerobic ribonucleoside-triphosphate reductase activating protein NrdG	-2.5
SPD0190	Hypothetical protein; predicted uridine kinase	-2.4
SPD0191	Anaerobic ribonucleoside-triphosphate reductase NrdD	-2.5
SPD0458	Heat-inducible transcription repressor HrcA	1.9
SPD0459	Heat shock protein GrpE	2.0
SPD0460	Chaperone protein DnaK (Heat shock protein 70)	1.7
SPD0461	Chaperone protein DnaJ	1.5
SPD0502	PTS system, β-glucosides-specific IIABC components	2.3
SPD0503	6-Phospho-β -glucosidase BglA-2	2.2
SPD0633	Copper transport repressor, CopY/TcrY family CopY	1.9
SPD0634	Copper -translocating P-type ATPase/cupredoxin CupA	2.2
SPD0635	Cation-transporting ATPase, E1-E2 family CopA	1.9
SPD0888	Adhesion lipoprotein AdcAII (LmB)	10.1
SPD0889	Pneumococcal histidine triad protein D precursor PhtD	12.0
SPD0890	Pneumococcal histidine triad protein E precursor PhtE	24.6
SPD0891	Hypothetical protein	3.8
SPD0893	Hypothetical protein	4.0
SPD0913	Inosine-5'-monophosphate dehydrogenase	3.4
SPD1038	Pneumococcal histidine triad protein A precursor PhtA	8.7
SPD1594	Transcriptional regulator	2.2
SPD1595	Hypothetical protein	2.1
SPD1636	Zinc-containing alcohol dehydrogenase AdhB	-1.6
SPD1637	Transcriptional regulator, MerR family	-1.8
SPD1638	Cation efflux system protein CzcD	-2.3
SPD1726	Pneumolysin PlY	-2.4
SPD1727	Hypothetical protein	-2.4
SPD1728	Hypothetical protein	-2.4
SPD1729	Hypothetical protein	-2.4
SPDI865	Zinc-containing alcohol dehydrogenase AdH	-2.9
SPD1986	L-fucose isomerase FucUL	-2.3
SPD1987	Fucolectin-related protein	-2.1
SPD1988	Glycoside hydrolase family 95	-2.0
SPD1989	PTS system, IID component	-3.0
SPD1990	PTS system, IIC component	-2.5
SPD1991	PTS system, IIB component	-2.6
SPD1992	PTS system, IIA component	-2.1
SPD1993	Fucose operon FucU protein/ RbsD/FucU transport protein FucUL	-1.9
SPD1994	L-Fuculose phosphate aldolase FucA	-2.9
SPD1995	Putative L-fuculose kinase FucK	-2.2
SPD1997	Zinc ABC transporter, zinc-binding lipoprotein AdcA	4.8
SPD1998	Zinc ABC transporter, permease protein AdcB	3.4
SPD1999	Zinc ABC transporter, ATP-binding protein AdcC	5.0
SPD2000	*adc* operon repressor AdcR	5.4
SPD2068	Serine protease	2.2
SPD2069	SpoJ protein	2.3

A



B

```
TACTTTACTGGTAAATA-N25-GCTTGAAGAAACAAAGAAAAAGGCTTACAATTGAACTATAAATATTACAAATAAAA-N17-ATG --- adh
ATACTTAAATGGTTCATAAAATTCTTGACAAGTTGGAGATTTAGGAGTAAACTATTAACCAGTTAAGTAATAGAGAGG-N59-ATG --- phtA
TTTTTAGAAAAACGTAACAGAAACTTGACAAAAGTAATTTTAAATAGTAAACTATTTACTGGTTAATTAAATGGTTAA-N78-ATG --- phtE
TTATTATGTATATATGTAAAATTCTTGACAAGCAATATTAAAAAGAGTAAACTATTAACTAGTTAATTAACCGGTTTA-N43-ATG --- phtB
TTTCAAATTTTTTCTAAAAATATCTTGACAGATTAAATTTTCAGGAGTAGAATATTTACTAGTTAATTAAATGGTTAA-N10-ATG_--- lmb
ATGAAACTAGTTTTCACGTTTTTTCTTGACTCTTTTTTTGCTTGTGATATAATTAACTGGTAAACAAAATGTCAAAGA-N50-ATG --- adcR
```

Figure 1: Organization of the *adcR* operon and the AdcR target genes *adcAII- phtD*, *phtA*, *phtB*, *phtE*, and *adh*. **A)** A model of Zn^{2+}-dependent gene repression by AdcR, **B)** Sequence of P*adcR*, P*adcAII*, P*phtA*, P*phtB*, P*phtE* and P*adh*. Predicted core promoter sequences are in bold while putative AdcR binding motifs are underlined. In some places, less relevant nucleotides are replaced with N and the number of replaced nucleotides.

Zn^{2+}-dependent expression of *adcRCBA, adcAII, adh* and the *pht* genes

In order to investigate in more detail the transcriptional regulation of genes that were found in the Zn^{2+} limitation microarray experiment, ectopic transcriptional *lacZ*-fusions were constructed to the *phtA, phtB, phtE, adcAII, adcR,* and *adh* promoters in the *bgaA* locus of D39 wild-type (88). The expression from the promoters of *adcR*, *phtA, phtB, phtE* and *adcAII* was 5-6 fold higher in CDM without supplemented Zn^{2+} as compared to CDM with 0.2 mM Zn^{2+} (Table 4). In case of *adh,* lower expression was observed during Zn^{2+}-limitation as compared to the high Zn^{2+} condition. This confirms the results of the Zn^{2+}-limitation microarray and demonstrates that the Zn^{2+} effects occur via the identified promoter regions.

Microarray analysis with the *adcR* mutant

In order to investigate in more detail the role of AdcR in the Zn^{2+}-dependent transcriptional effects, we replaced the *adcR* gene with an erythromycin-resistance (ery^R) cassette by allelic replacement mutagenesis. To analyze the effect of the *adcR* mutation on the

transcriptome of *S. pneumoniae*, the wild-type strain D39 was compared with its isogenic *adcR* mutant by using DNA microarrays. As AdcR was expected to be active in the presence of Zn^{2+}, we grew the cells in CDM supplemented with 0.2 mM Zn^{2+} up to mid-exponential phase. 24 genes were differentially expressed, of which 17 were up and 7 were down (Table 5).

Table 4: Specific β-galactosidase activity in miller units of the wild-type D39 and *ΔadcR* strains containing the *lacZ*-transcriptional fusions to P*adcR*, P*adcAII*, P*phtA,* P*phtB*, P*phtE* and P*adh*, grown in CDM with or without 0.2 mM Zn^{2+}. Standard deviations are given between parentheses.

Specific β-galactosidase Activity (Miller Units)				
Promoter	**D39 (wt)**		***ΔadcR***	
	0 mM	**0.2 mM**	**0 mM**	**0.2 mM**
adcR	250 (29)	55 (7)	745 (89)	805 (98)
adcAII	197 (25)	46 (9)	1225 (93)	1195 (108)
phtA	310 (34)	42 (5)	1450 (135)	1323 (127)
phtB	237 (17)	57 (11)	1170 (97)	1210 (121)
phtE	255 (12)	37 (6)	2010 (105)	2200 (174)
adh	26 (4)	159 (7)	8 (1)	6 (3)

The *adh* gene was highly downregulated in the *adcR* mutant, strengthening the hypothesis that it is activated by AdcR. Moreover, the *adc* operon was highly downregulated in the *adcR* mutant. The downregulation of *adcABC* is likely due to the polar effect of the mutation in *adcR*, as the *adcR* gene was replaced with an erythromycin-resistance cassette. Highly upregulated were the genes encoding the members of the Pht family of proteins (PhtA, PhtD and PhtE) and AdcAII. Therefore, it is likely that AdcR acts as a Zn^{2+}-dependent repressor of these genes. A cellobiose utilization operon was also significantly upregulated, although its relation to AdcR is unclear. In conclusion, the two DNA microarray experiments performed in this study indicate that AdcR mediates the Zn^{2+}-dependent regulation of its own operon, *adcAII- phtD*, *phtA*, *phtE* and *adh*.

AdcR regulates the expression of *phtA, phtB, phtE, adcAII, adh* and *adcR*

To demonstrate if the observed regulation via AdcR occurs through the promoter regions of the identified genes, transcriptional *lacZ* -fusions to the promoters of *phtA, phtB, phtE, adcAII* and *adcR* were introduced in the *adcR* mutant. Measurement of specific β-galactosidase activity (Miller Units) showed high expression of the *phtA, phtB, phtE, adcAII,* and *adcR* promoters even in the presence of Zn^{2+} (Table 4). Therefore, AdcR acts as Zn^{2+}-

dependent repressor of the *adcRCBA*, *adcAII-phtD*, *phtA*, *phtB* and *phtE* operons, by regulating their respective promoters.

Table 5: Summary of transcriptome comparison of *S. pneumoniae* wild-type strain with SS200 grown in CDM with 0.2 mM Zn^{2+} [a]Gene numbers refer to D39 locus tags. [b]D39 annotation/ TIGR4 annotation (100, 136, 243), [c]Ratios >2.0 or <-2.0 (Signal intensity for SS200 divided by that for wild-type strain.

Gene tag[a]	Function[b]	Ratio[c]
SPD0277	6-Phospho-β -glucosidase	10.6
SPD0278	Hypothetical protein	2.7
SPD0279	PTS system, IIB component	9.4
SPD0280	DNA Binding transcriptional regulator	8.3
SPD0281	PTS system, IIA component	6.1
SPD0282	Hypothetical protein	6.8
SPD0283	PTS system, IIC component	6.2
SPD0308	ATP-dependent Clp protease, ATP-binding subunit, putative	2.3
SPD0442	Integrase/recombinase, phage integrase family	3.2
SPD0888	Adhesion lipoprotein AdcAII (LmB)	11.5
SPD0889	Pneumococcal histidine triad protein D precursor PhtD	14.0
SPD0890	Pneumococcal histidine triad protein E precursor PhtE	47.6
SPD0892	Hypothetical protein	5.6
SPD0893	Hypothetical protein	33.6
SPD1038	Pneumococcal histidine triad protein A precursor PhtA	77.2
SPD2018	Hypothetical protein	2.9
SPD0450	Type I restriction-modification system, S subunit, putative	-3.1
*SP0763**	Hypothetical protein	-2.2
SPD1865	Zinc-containing alcohol dehydrogenase AdH	-13.2
SPD1997	Zinc ABC transporter, zinc-binding lipoprotein AdcA	-7.7
SPD1998	Zinc ABC transporter, permease protein AdcB	-4.1
SPD1999	Zinc ABC transporter, ATP-binding protein AdcC	-5.2
SPD2000	*adc* operon repressor AdcR	-7.4

*33% identity with *SPD1458*

To investigate further the role of AdcR as an activator, the transcriptional *lacZ*-fusion to P*adh* was transferred to the *adcR* mutant. The *adcR* mutation led to total inactivation of the *adh* promoter even in the presence of 0.2 mM Zn^{2+} (Table 4). This shows that AdcR not only acts as a Zn^{2+}-dependent repressor but also as a Zn^{2+}-dependent activator.

Binding of AdcR to its targets is Zn^{2+}-dependent

To investigate whether the *adcR* targeted promoters, as found in the transcriptional analyses, are under the direct control of the AdcR protein, EMSAs were performed with purified strep-tagged AdcR (Ad-strep) and ^{33}P labeled promoter regions of *adcR*, *adcAII*, *phtA*, *phtB, phtE* and *adh*. The promoter region of *SPD1829,* which encodes a putative transcriptional regulator belonging to the GntR family, was taken as a negative control, since in the high and low Zn^{2+} transcriptome data, expression of *SPD1829* was unaffected. In the absence of Zn^{2+}, Ad-strep was not able to efficiently shift the promoter regions of *adcR*,

adcAII, *phtA*, *phtB, phtE*, and *adh.* Only at a high concentration, Ad-strep was able to bind (Fig. 2-II). As we have excluded EDTA from all the solutions used for Ad-strep protein purification, this might be due to the presence of trace amounts of Zn^{2+} in the purified Ad-strep samples, a phenomenon that was also seen in a previous study on a Zn^{2+} binding regulator (128). Indeed, no shift was observed even at the highest Ad-strep concentration in the presence of EDTA (Fig-2-I). In the presence of 0.2 mM Zn^{2+} Ad-strep already bound to the promoter regions of *adcR*, *adcAII*, *phtA*, *phtB, phtE,* and *adh* genes even when present in a low concentration (Fig. 2-III). Under the conditions tested, no shift was observed in the negative control (Fig. 2A). Therefore, our data demonstrate unequivocally that AdcR interacts directly with its targets in the presence of Zn^{2+}.

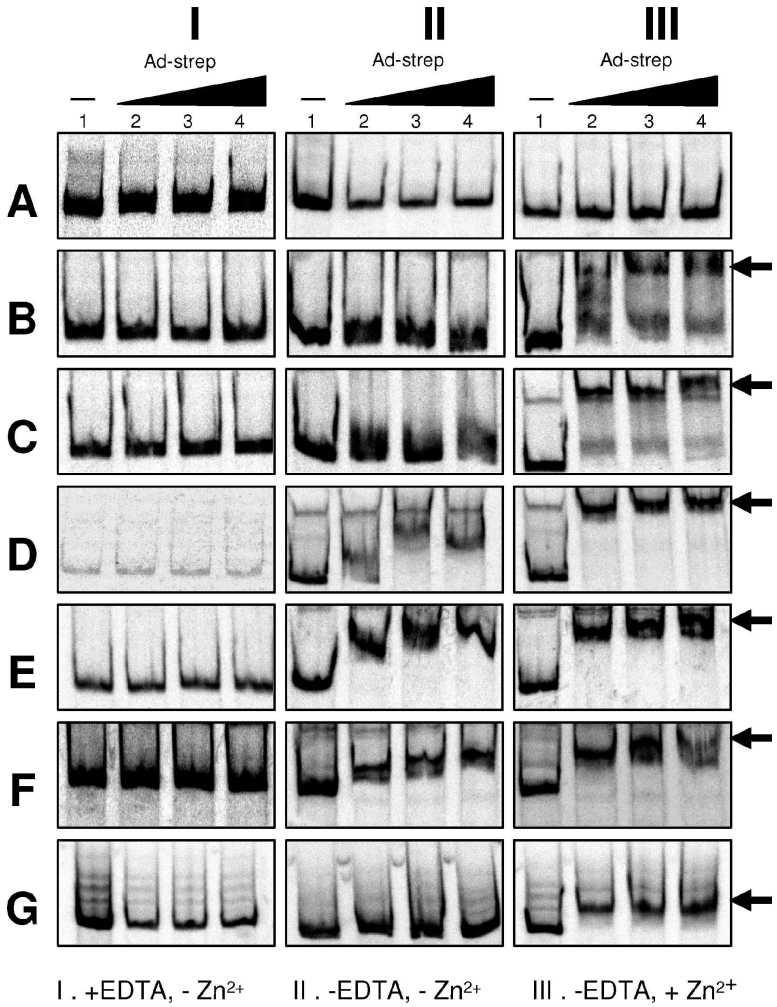


Figure 2: *In vitro* interaction of Ad-strep with the promoter regions of *SPD1829* **(A)**, *adcR* **(B)**, *adcAII* **(C)**, *phtE* **(D)**, *phtA* **(E)**, *adh* **(F)** and *phtB* **(G)**. **I:** with EDTA (1 mM) and without 0.2 mM Zn^{2+}. **II:** without EDTA and without Zn^{2+}. **III:** without EDTA (1 mM) and with 0.2 mM Zn^{2+}. Purified Ad-strep was added at three different concentrations (270, 540 and 800 nM) from lane 2 to 4 while lane 1 is without protein. The triangular bar above lanes 2 to 4 indicates the presence of Ad-strep. Arrows indicate the position of the shifted probes. The presence of weaker bands that run higher than the free probe in the gels is a phenomenon that has also been seen by others in similar experiments. These bands may represent unspecific PCR products or single-stranded DNA (4, 66).

To investigate whether the Ad-strep-DNA interaction is Zn^{2+}-specific or whether other trace metal ions also have a role, EMSAs was performed with selected promoters in the presence of metal ions like Mn^{2+}, Fe^{2+}, Fe^{3+}, Ca^{2+}, Co^{2+}, Ni^{2+} and Cu^{2+} in a concentration of 0.2 mM (Fig. 3). The data showed that, except Zn^{2+}, none of the tested metals were able to shift the P*phtB* (a repressed promoter) (Fig. 3A) and P*adh* (an activated promoter) (Fig. 3B) region. Therefore, the interaction of AdcR with its targeted promoters (*adcR*, *adcAII*, *phtA*, *phtB*, *phtE* and *adh*) is highly Zn^{2+}-specific.

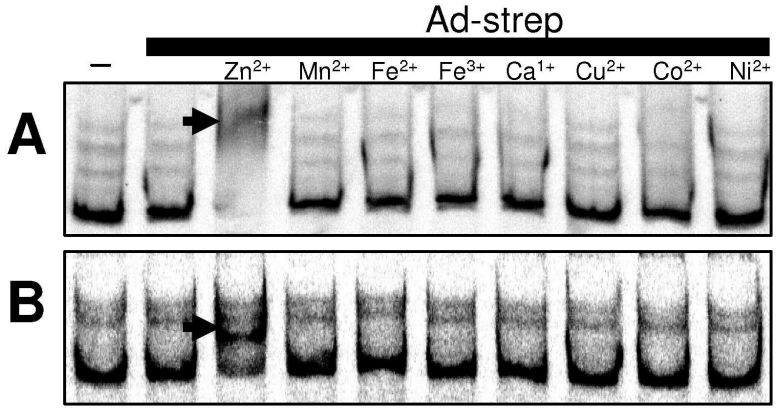


Figure 3: Metal dependence of the *in vitro* interaction of Ad-strep with the *phtB* **(A)** and *adh* **(B)** promoter regions. The horizontal bar above lanes 2 to 8 indicates the presence of 270 nM Ad-strep. Metal ions were added as indicated above the lanes in a concentration of 0.2 mM. Lane 1, free probe. Arrows indicate the position of the shifted probes.

Discussion

Previous studies indicate that Zn^{2+} is an important trace metal ion in the life-style of Gram-positive pathogens, like *S. pneumoniae*. Recently, the expression of several virulence genes, encoding for example the choline binding protein PcpA, the extracellular protease PrtA and the lipoprotein Mn^{2+}-uptake system PsaBCA, was shown to be strongly induced in *S. pneumoniae* by a high concentration of Zn^{2+} (129). In this study, we have analyzed the transcriptome change in pneumococcus in response to Zn^{2+} limitation. Our transcriptional analysis shows that the expression of various genes and operons of diverse functions, including the AdcR regulon, is altered during Zn^{2+} limitation in *S. pneumoniae*. Thus, the concentration of Zn^{2+} has a quite significant impact on the transcriptome of *S. pneumoniae*. No downregulation of *pcpA*, *prtA* and *psaBCA* was seen, likely because the medium contained 7 µM Mn^{2+}, which is high enough to repress expression of these genes (129). In subsequent experiments, the main part of the transcriptional effects in Zn^{2+} limitation was shown to be mediated by the Zn^{2+}-dependent regulator AdcR.

Determination of the intracellular concentration of trace metal ions in cells grown in CDM with a high and low content of Zn^{2+} showed that only the levels of Zn^{2+} were changed. Since, the cell-associated concentration of Cu, Co, Ni and Fe was around or lower than the detection limit, we cannot exclude that differences in the concentration of these metal ions are also partly responsible for the effects seen in the transcriptome analysis on Zn^{2+} limitation. For example, upregulation of the *cop* operon could indicate that the level of Cu was higher in the Zn^{2+}-limited condition. However, Zn^{2+} is also known to repress this operon, which could explain this effect as well (225). The results of the *lacZ*-reportor studies and EMSAs strongly support the conclusion that most responses observed in the transcriptome study are due to Zn^{2+} limitation, and are mediated by AdcR. In bacteria, most Zn^{2+} is probably bound to proteins (187) like ribosomal proteins, metal-dependent regulators and other metalloproteins (92, 188). Therefore, it is very likely that in Zn^{2+} limitation less Zn^{2+} is available for binding to AdcR, which in its apo-form does not bind to its target promoters, thereby causing de-repression of the downstream genes

In *Escherichia coli*, the level of free Zn^{2+} is tightly controlled by the Zn^{2+}-responsive regulators ZuR (Zinc Uptake Regulator) and ZntT, which control Zn^{2+}-uptake and -export, respectively (174). In this bacterium, Zn^{2+}-depletion results in the upregulation of 9 genes including *zinT* (*yodA*) and *ZunA* (83). *ZinT* is regulated by ZuR and has homology with the *S. pneumoniae adcA* gene, while *zunA* has homology with *adcAII* and *adcA*. Expression of other Zn^{2+}-dependent genes including *znuBC*, which has homology with *adcCB* and *psaBC* in *S. pneumoniae*, was also altered in *E. coli*, but their ratio was lower than the cut-off value of 2. In *Bacillus subtilis* Zur is found to control the Zn^{2+}-starvation response. In response to a high Zn^{2+} concentration, ZuR causes the repression of two Zn^{2+}-uptake systems: The *ycdHI-yceA* encoded ABC transporter is the primary Zn^{2+}-uptake system while the *yciABC* operon may encode a second, lower affinity system (174). The *ycdHI-yceA* operon has homology with the *adcCBA* genes in *S. pneumoniae*, but no homolog has been found in *S. pneumoniae* for ZuR and YciABC. *B. subtilis* has a third Zn^{2+}-uptake system, which is ZosA (174). ZosA has homology to ZntA of *E. coli* and *pmtA* (PerR regulated metal transporter A), the last one which was found to be upregulated in the *perR* mutant microarray in *Streptococcus pyogenes* (34). Interestingly, overexpression of *pmtA* by deletion of *perR* in *S. pyogenes* led to upregulation of the AdcR regulon, and downregulation of the Zn^{2+}-resistance gene *czcD*, indicating a link between oxidative-stress resistance and Zn^{2+} regulation (34). In *S. pyogenes,* the deletion of AdcR also causes upregulation of *rpsN2*, a 30S ribosomal protein that contains a Zn^{2+}-binding CXXC motif (34). *S. pneumoniae* has a gene homologous to *rpsN2* (*rpsN*), but

no change in the expression of *rpsN* was observed in both our transcriptomes. In *S. pneumoniae*, the gene with the highest homology with *S. pyogenes pmtA* is *SPD0635*, encoding a copper influx protein CopA, which is part of an operon with *copY* and *cupA*. Interestingly, the *cop* operon in *S. pneumoniae* was upregulated during Zn^{2+} limitation, suggesting a link between Zn^{2+} and Cu^{2+} homeostasis. In *S. suis* the AdcR protein was found to be responsible for the regulation of the Pht proteins in response to Zn^{2+} (12). In conclusion, in the above organisms Zn^{2+}-transport systems are obviously regulated by Zn^{2+}, but in streptococci, including *S. pneumoniae*, a number of putative virulence genes, most notably the Pht family proteins, are affected by Zn^{2+} as well.

AdcR belongs to GntR family regulators and its mechanism of Zn^{2+}-responsive repression via AdcR has previously been studied in other streptococci (11, 12, 34, 151, 173). Here we have shown by detailed genetic and biochemical analyses that AdcR acts both as Zn^{2+}-responsive repressor as well as a Zn^{2+}-responsive activator in *S. pneumoniae*, by regulating directly the promoters of its target genes. Several promoters of the genes found in Zn^{2+}-limitation microarray have a putative AdcR operator site (5'-TTAACYRGTTAA-3') in their promoter regions (10, 188). *In silico* analysis showed that the position of AdcR binding motif is important, as all the AdcR repressed genes have AdcR operator site in upstream of the core promoter, while in case of activation an operator site is found in the downstream of the core promoter. Since Reyes-Caballero *et al*. (207) showed that the AdcR binding site is functional at least in the *adcR* promoter, it is highly likely that AdcR exerts its function by binding directly to the predicted operators in the promoters of the operons that were analyzed in this study.

In this study, we show for the first time, that AdcR interacts in a Zn^{2+}-dependent manner with its target promoters. Our data conflict with the data of Ogunniyi *et al.* (185), who showed that AdcR only interacts with the *phT* promoters in the absence of Zn^{2+}. However, in agreement with our data, Reyes-Caballero *et al.* (207) have shown that AdcR interacts with the *adcR* promoter only in the presence of Zn^{2+}. Also, in *Streptococcus suis* AdcR binds to the promoters of the *pht* genes in the presence of Zn^{2+}. Because of these reasons, and because both our transcriptional data and our biochemical data are in perfect agreement, we believe AdcR is actively binding to DNA only in the presence of Zn^{2+}, not in its absence. Reyes-Caballero *et al.* (207) also saw that Co^{2+} and Mn^{2+} activate AdcR DNA binding activity, and in *S. suis* Mn^{2+} stimulated the AdcR- interaction with its operator sequence in its targets (10). A similar effect could however not be observed in this study. Also, adding Co^{2+}, Mn^{2+}, Fe^{2+}

and Ni^{2+} to the medium did not lead to activation of the *adh* promoter, like was seen with Zn^{2+} (data not shown). Therefore, AdcR seems highly specific for Zn^{2+} in *S. pneumoniae*.

Because it regulates the *pht* and *adcAII* genes, AdcR has a possible impact on the virulence of *S. pneumoniae*. Pneumococcal histidine triad (Pht) family proteins are characterized by a HxxHxH motif that usually occurs multiple times throughout the protein sequence. These proteins were only found in pathogenic streptococcal species and involved in the invasion process (2, 188). *S. pneumoniae* has four annotated Pht family proteins (PhtA, PhtE, PhtB and PhtD). It has recently been shown that humans produce antibodies to Pht proteins upon exposure to pneumococcus, and that immunization of mice has provided protective immunity against sepsis and pneumonia as well as reduced nasopharyngeal colonization (12, 169). Although the contribution of the Pht proteins in the virulence of *S. pneumoniae* is poorly understood, some recent studies showed that PhtA, PhtB, PhtD and PhtE can act as protective antigens in multicomponent pneumococcal protein vaccines in mice (169). Moreover, the laminin-binding protein (Lmb or AdcAII) was 10-fold upregulated in the Zn^{2+}-limitation transcriptome and is involved in Zn^{2+}-acquisition (149). Recombinant laminin binding protein of *S. pyogenes* (LbP) or *S. agalactiae* (LmB) was able to bind human laminin in an *in vitro* experiment (72, 235, 239, 240), while its deletion in *S. mutans* resulted in decreased eukaryotic cell attachment and internalization (72, 235, 239). Therefore, it is likely that Zn^{2+} affects the virulence status of pneumococcus, by changing the expression of several pathogenicity factors.

Conclusion

This study complements our previous study, where we identified the pneumococcal virulence genes that are induced in response to a high concentration of Zn^{2+}. Here, we prove that a significant number of genes is affected by Zn^{2+} limitation, amongst which the putative virulence genes *phtA*, *phtB*, *phtD*, *phtE*, and *adcAII*. All of these genes were shown to belong to the regulon of AdcR, which was demonstrated to act both as a transcriptional repressor and activator in the presence of Zn^{2+}. The data imply that the virulence status of *S. pneumoniae* is significantly affected by fluctuating Zn^{2+} concentrations.

Acknowledgements

We thank Anne de Jong and Siger Holsappel for help with the DNA microarray slide production. Funds for SS were made available under the HEC-NUFFIC overseas scholarship

program by HEC (Higher Education commission) Pakistan. TGK is supported by the EU-funded Pneumopath project HEALTH-F3-2009-222983.

Chapter 3

The *cop* operon is required for copper homeostasis and contributes to virulence in *Streptococcus pneumoniae.*

Sulman Shafeeq, Hasan Yesilkaya, Tomas G. Kloosterman, Geetha Narayanan, Michal Wandel, Peter W. Andrew, Oscar P. Kuipers and Julie A. Morrissey

Based on *Mol. Microbiol*. **81**: 1255-70 (2011).

Abstract

High levels of copper are toxic and therefore bacteria must limit free intracellular levels to prevent cellular damage. In this study, we show that a number of pneumococcal genes are differentially regulated by copper, including an operon encoding a CopY regulator, a protein of unknown function (CupA) and a P1-type ATPase, CopA, which is conserved in all sequenced *Streptococcus pneumoniae* strains. Transcriptional analysis demonstrated that the *cop* operon is induced by copper *in vitro*, repressed by the addition of zinc and is auto-regulated by the copper-responsive CopY repressor protein. We also demonstrate that the CopA ATPase is a major pneumococcal copper resistance mechanism and provide the first evidence that the CupA protein plays a role in copper resistance. Our results also show that copper homeostasis is important for pneumococcal virulence as the expression of the *cop* operon is induced in the lungs and nasopharynx of intranasally infected mice, and a *copA*$^-$ mutant strain, which had decreased growth in high levels of copper *in vitro*, showed reduced virulence in a mouse model of pneumococcal pneumonia. Furthermore, using the *copA*$^-$ mutant we observed for the first time in any bacteria that copper homeostasis also appears to be required for survival in the nasopharynx.

Introduction

Streptococcus pneumoniae is the main cause of bacterial pneumonia worldwide and is also a major agent of otitis media, bacteraemia, and meningitis. For *S. pneumoniae* to survive and cause infection in the very different host environments associated with these diseases, it must be able to sense and adapt to considerable variation in environmental conditions, including changes in the concentration of metal ions in host tissues.

Transition metal ions are essential co-factors for many enzymes but they can also be highly toxic. Therefore, bacterial metal ion homeostasis is extremely important to ensure sufficient intracellular levels of metal ions for use as co-factors, but also to limit excess intracellular levels to prevent toxicity. Several metal ion transport systems have been implicated in pneumococcal virulence (35, 95, 165, 212), and furthermore the concentration of metal ions can influence virulence gene expression (86, 110, 128, 224). Under conditions of limitation, metal ions are imported into *S. pneumoniae* by specific metal ion transporters such as PsaBCA, PitABCD and AdcCBA, which are responsible for the transport of manganese, iron and zinc respectively (69, 70, 111, 212, 224), while an excess of metal ions

is removed from the cell by specific efflux systems, for example the zinc exporter CzcD (128) and the manganese exporter MntE (107, 212).

Copper is also an important transition metal for most organisms, albeit toxic at high levels. So far, copper homeostasis has not been studied in *S. pneumoniae*. Like in many Gram-positive bacteria, known copper-containing proteins have not been identified in the *S. pneumoniae* genome sequences (232). However, there is some evidence that pneumococci may require copper, as a *S. pneumoniae* strain mutant for the (p)ppGpp synthetase *relA* is unable to grow in chemically defined medium (CDM) unless it is supplemented with copper and manganese (116). Even if copper is not used, pneumococci must still have mechanisms to export an excess of copper from the cell, as *S. pneumoniae* will encounter varying levels of copper *in vivo.* In some tissues, especially in the blood, free copper levels are very low, but levels can be higher in tissues, for example the lungs (lungs 121.96 and blood 12.98 μg/g dry weight, (46)). Moreover, copper levels in the serum have been shown to increase during infection (13).

To maintain copper homeostasis and prevent toxicity, bacteria use a number of efflux and sequestration mechanisms to remove excess copper, and also initiate a global adaptive genetic response which can involve induction of other stress regulons (18, 118, 238, 261). However, the cause of toxicity, the mechanisms of resistance and the regulatory responses used can vary significantly between species. Resistance mechanisms include efflux systems, such as the ubiquitous CopA/CopB P1-type ATPase transporters (257), and sequestration mechanisms, including the CopZ family of copper binding proteins, which chaperone the copper ions intracellularly for incorporation/use/efflux by other copper-binding proteins (197). Two main types of Gram-positive copper-responsive regulators have been identified to date. These are the CopY copper-responsive repressor family found in *Enterococcus* and *Streptococcus* spp. (197), and the CsoR copper-responsive repressors found in *Mycobacterium tuberculosis* (146) and *Bacillus subtilis* (231).

The objective of this study was to investigate the mechanism of pneumococcal copper homeostasis and its role in virulence. Transcriptional profiling with DNA microarrays identified a number of genes that are differentially expressed depending on the copper concentration, including a conserved tri-partite operon encoding a homologue of the CopY regulator family (*copY*), a hypothetical protein (*cupA*) and a P1-type ATPase (*copA).* Transcriptional analysis by real time quantitative RT- PCR and assaying strains carrying transcriptional *lacZ*-fusions demonstrated that expression of the *cop* operon is induced specifically by copper *in vitro* and is auto-regulated by CopY. Expression of the *cop* operon,

as well as several other putative copper transport genes, is also induced in pneumococci isolated from the lungs and nasopharynx of intranasally infected mice. Furthermore a *copA*$^-$ mutant strain showed decreased virulence in a mouse model of pneumococcal pneumonia and a decreased ability to survive in the mouse nasopharynx, showing that copper homeostasis plays an important role in *S. pneumoniae* physiology and virulence.

Material and Methods

Bacterial strains and growth conditions

Table 1: Bacterial strains and plasmids used in this study.

Strain/plasmid	Description	Source
S. pneumoniae		
D39	Serotype 2 strain, *cps 2*	Laboratory of P. Hermans
P*copY-lacZ*	D39 Δ*bgaA*::P*spd0633-lacZ*; Tet^R	This study
P*copY-wt-lacZ*	D39 Δ*bgaA*::P*spd0633-C-lacZ*; Tet^R	This study
P*copY-mut-lacZ*	D39 Δ*bgaA*::P*spd0633-F-lacZ*; Tet^R	This study
copY-stop	D39 *copY* null mutant	This study
copY-stop(PcopY-lacZ)	D39 *copY;* Δ*bgaA*::P*copY-lacZ*; Tet^R	This study
copY-stop-com	D39 *copY;* Δ*bgaA*:: P*copY+copY-lacZ*; Tet^R	This study
cupA-stop	D39 *cupA* null mutant	This study
Δ*copA*	D39 *spd0635*::Spec^R	This study
Δ*cupA copA*	D39 *spd0634-5*::Spec^R	This study
copA$^-$	D39 *spd0635*::Spec^R	This study
E. coli		
EC1000	Km^R; MC1000 derivative carrying a single copy of the pWV1 *repA* gene in *glgB*	Laboratory collection
Plasmids		
pPP2	Amp^R Tet^R; promoter-less *lacZ.* For replacement of *bgaA* with promoter *lacZ*-fusion. Derivative of pTP1	(88)
pORI280	Erm^R; *ori*$^+$ *repA*$^-$*;* deletion derivative of pWV01; constitutive *lacZ* expression from P32 promoter	(141)
pSS1	pORI280 *copY-stop,* containing two stop mutations after 6 base of coding sequence	This study
pSS2	pORI280 *cupA-stop,* containing two stop mutations after 21 base of coding sequence	This study
pSS3	pPP2 P*copY-lacZ,* containing the whole promoter	This study
pSS4	pPP2 P*copY-wt-lacZ,* containing *cop* motif 1 and 2	This study
pSS5	pPP2 P*copY-mut-lacZ* containing *cop* motif 1 only	This study
pSS6	pPP2 P*copY+copY-lacZ*	This study

Bacterial strains used for this study are listed in Table 1. Pneumococci were grown as static cultures in M17 (110) broth containing 0.5% (w/v) glucose (GM17) or chemically defined medium (CDM) containing 0.5% (w/v) glucose at 37 °C in air or in brain heart

infusion (BHI) broth or on blood agar plates supplemented with 1% (v/v) defibrinated sheep blood or 5% (v/v) defibrinated horse blood, in microaerophilic conditions at 37°C. CDM was prepared as described before (125), with the exception that $ZnSO_4$ and $CuSO_4$ were omitted from the metal mixture and added separately as specified in the Results section. Metal ions were added as the salts $ZnSO_4$, $MnSO_4$, $MgCl_2$, $CaCl_2$, $CoCl_2$, $NiSO_4$, $CuSO_4$ and $FeCl_2$. For growth experiments, β-galactosidase assays and transcriptome analysis, *S. pneumoniae* cells frozen at an optical density at 595 nm of 0.3 in GM17/CDM were washed once with the appropriate medium and diluted 1:40 for inoculation. To determine the copper susceptibility of the *copA*$^-$ mutant strain, Sicard's defined medium (229) was supplemented with 0 to 2.5 mM Cu^{2+} and the increase in optical density at 500 nm was determined every hour over 16 h. *Escherichia coli* strain EC1000 was cultured at 37 °C on TY agar or TY broth. Where necessary for selection, media were supplemented with the following concentrations of antibiotics: erythromycin: 0.25 μg ml^{-1} for *S. pneumoniae* and 120 μg ml^{-1} for *E. coli*; spectinomycin: 100 μg ml^{-1} and tetracycline: 2.5 μg ml^{-1} for *S. pneumoniae*; and ampicillin: 100 μg ml^{-1} for *E. coli*. When appropriate, 0.006% (w/v) X-Gal was used in plates. Long-term storage of bacteria was done at -80 °C in 10% (v/v) glycerol.

DNA isolation and manipulation

Primers used in this study are listed in Table 2. Chromosomal DNA of *S. pneumoniae* D39 wild-type strain was used as a template for PCR amplification (9,59). All DNA manipulations were done as described before (125).

Construction of mutants

The *copY-stop* and *cupA-stop* mutants were constructed using plasmid pORI280 as described before (126). Briefly, primers SPD0633-2 NcoI/SPD0633-3 NcoI and SPD0634-2 NcoI/SPD0634-3 NcoI, which introduce four premature stop codons and a NcoI site in frame after the first 6 bases in the *copY* (SPD0663) and 21 bases in the *cupA* (SPD0634) open reading frame, were used in combination with primers SPD0633-1 XbaI and SPD0633-4 EcoRI and SPD0633-1 BglII and SPD0633-4 XbaI respectively, to amplify fragments comprising the upstream and downstream sequence of *copY* (SPD0633) and *cupA* (SPD0634). These PCR products were ligated using the NcoI site. The resulting products were cloned respectively as an XbaI, EcoRI, or BglII, XbaI fragment in pORI280, giving plasmids pSS1 and pSS2. The mutations led to the appearance of an NcoI site, on the basis of which the desired mutant could be identified. The mutations were further verified by DNA sequencing.

pSS1 and pSS2 were used to introduce the mutations into the chromosome of *S. pneumoniae* D39 as described before (125), giving the *copY-stop* and *cupA-stop* mutant strains.

Table 2: Primers used in this study.

Name	Nucleotide sequence (5' to 3')[a]	Restriction site
FcopY	CGGAATTCTTACGCACCCCTAGTGGTTG	EcoRI
RcopY	CGGGATCCGACCTGCCATTCTGCATCTG	BamHI
PcopY-C	CATGGGATCCTCCAAAATCTACATTTGTC	BamHI
PcopY-F	CATGGGATCCACACTCATCAGTATACTC	BamHI
CopY-Com	CGGGATCCGTCTCCTTTACATACAATTAC	BamHI
Spd0633-1 XbaI	GCTCTAGACCACTTGTAGCAACCATAAC	XbaI
Spd 0633-2 NcoI	CCGCCATGGTTATTACTGCATTACCATACCTCC	NcoI
Spd 0633-3 NcoI	CGCCCATGGTAATAAATTTCAGATGCAGAATGGC	NcoI
Spd 0633-4 EcoRI	GCTGGAATTCCCTTTTGAAGCTGG	EcoRI
Spd-0634-1-BglII	GAAGATCTGCAATGGTCAAAGCAACGG	BglII
Spd-0634-2-NcoI	CATCCATGGTTATTAGGTTACAATACTATTTAACATGAC	NcoI
Spd-0634-3-NcoI	CATCCATGGTAATAAATTATTTGTATTGCCCTTATCG	NcoI
Spd-0634-4-XbaI	TGCTCTAGAATCACAAGAGGCACAGTAAAG	XbaI
Spd 0634-1	CCTGCAATGGTCAAAGCAACGG	-
Spd 0634-2	GCATGGCGCGCCGACGTCTCCTTTACATAC	AscI
Spd 0635-1	GAAGATTATTTGGATGCAGG	-
Spd 0635-2	GCATGGCGCGCCAGTCTCCACCTACTCTAC	AscI
Spd 0635-3	GCATGCGGCCGCCTGTTCTGACCAGTTCCCTTCTG	NotI
Spd 0635-4	GGGCGTTCAGCATTGTTCAAG	-
Spec-F	GCATAGGCGCGCCCTAATCAAAATAGTGAGGAGGATAT	AscI
Spec-R	CGATTGCGGCCGCACTAAACGAAATAAACGCTAAAACG	NotI
RT-PCR primers		
Spd0633F	ACACTTTTGTCTCGGCTGGT	
Spd063R	TCCTACGGGAACAAACCTTG	
Spd0635F	TTCGTTCTTTTGGGAGCAGT	
Spd0635R	TCAGGTCGCCTACCTTGACT	
Spd0709F	TCGTGTGGCTGCCAAGCGTG	
Spd0709R	GGCTGATCCACCAGCTGAGTC	
Spd1436F	GGTCTTGGTGCAGGAGATGT	
Spd1436R	GCTGGCAATAGCCTCTTCAC	
Spd1920F	TAATGCCATCGTCAAGCAAG	
Spd1920R	TGCTCATCCTGCATCAAGAC	
Spd1118F	TACCCCCACCTGGTCAATAA	
Spd1118R	TCCTAACTCGTGCTGGTGTG	
MP127	CCGGGGACTTATCAGCCAACC	
MP128	TACTAGCGACGCCATCTATGTG	
CopY-F	GGAAGCTGTGATTTCAGAG	
CupA-R	GTATCCGTTTTTTTGCTGGG	
CupA-F	GAACAGGCTGGAGAGTTTAG	
CopA-R	CCTTTTGAAGCTGGATATGGG	

[a]Restriction enzyme sites are underlined.

Deletion strains of *copA* and *cupA-copA* were made with allelic replacement with a spectinomycin marker. Briefly, primers SPD0634-1/SPD0634-2, SPD0635-1/ SPD0635-2, SPD0635-3/SPD0635-4 were used to generate the PCR fragments of the left and right flanking regions of *copA* and *cupA-copA*, respectively. Then these PCR products were ligated using AscI/NotI site with the spectinomycin PCR product, which was generated with primers

Spec-F/Spec-R. The ligated product was transferred to *S. pneumoniae* D39. Spectinomycin resistance clones were examined for the presence of the *copA* or *cupA-copA* deletion by PCR.

To construct the *copA*$^-$ insertion mutant, the chromosomal region encompassing *copA* (SPD0635) was amplified with copAF and copAR primers (Table 2). The amplicons were incubated with *Himar1* transposase (135) and plasmid pR412, which contains the *mariner* mini-transposon conferring spectinomycin resistance (162). Then the *in vitro* mutagenised DNA was transformed into the pneumococcus (5). Pneumococcal transformants were selected for spectinomycin resistance, and insertion of the resistance cassette into the chromosome was confirmed by PCR by using transposon-specific primers, MP127 or MP128, with appropriate chromosomal primers, and sequenced as described previously (274). A representative strain, designated as *copA*$^-$, was selected for further study.

Complementation of *copY*

To complement the *copY* gene in the D39 *copY*-stop mutant, we amplified the native *copY* promoter and gene with primer pair FcopY and PcopY-com. This PCR product was further cloned into pPP2 plasmid (88) resulting in a plasmid pSS4. pSS4 was transferred to D39 *copY*-stop mutant strain to complement the *copY* gene.

Construction of *lacZ*-fusions

Chromosomal transcriptional *lacZ*-fusions to the P*copY* (whole *copY* promoter region), P*copY-F* (*cop* box 1 only), and P*copY-C* (*cop* box 1 and 2), were constructed in the integration plasmid pPP2 (88) via double crossover in the *bgaA* gene with primer pairs FcopY/ RcopY, FcopY/PcopY-F and FcopY/PcopY-C leading to plasmids pSS3-5 respectively. The *lacZ*-fusion constructs were introduced into wild-type D39, as well as the D39 *copY*-stop mutant as described above. All plasmid constructs were checked by sequencing, and new loci created with these plasmids were verified by PCR.

β-galactosidase assays

Specific β-galactosidase activity was measured as described before (93) using cells grown in CDM at 37 °C supplemented with different concentrations of metal ions as mentioned in the result section and harvested in the mid-exponential phase of growth.

Transcriptome analysis using *S. pneumoniae* DNA microarrays

DNA microarrays were used to determine Cu^{2+}-dependent pneumococcal gene expression (126). The expression in D39 strain grown in CDM supplemented with 0.05 mM $CuSO_4$ was compared to the transcriptome of the same strain in the absence of Cu^{2+}. The experiments were repeated with four biological replicates essentially as described previously (126, 253). In short, cultures were harvested at an optical density (OD) at 595 nm of approximately 0.3 by centrifugation for 1 min at 10000 rpm at room temperature. Cell pellets from 50 ml culture for each replication were immediately frozen in liquid nitrogen and store at -80 °C. RNA was isolated with the Roche RNA isolation kit. Synthesis of cDNA and Cy3/Cy5 labelling of 15-20 µg total RNA was performed with the CyScribe Post Labelling Kit (Amersham Bioscience). Hybridization was performed with labeled cDNA for 16h at 45 °C in Ambion Slidehyb #1 hybridization buffer on super-amine glass slides (Array-It, SMMBC). Slides were scanned with a Genepix 4200 laser scanner at 10 µm resolution. Array Pro 4.5 (Media Cybernetics Inc., Silver Spring, MD) was used to analyze the slides. The *MicroPrep* software package was used to obtain the microarray data from the slides. The expression ratio of D39 strain + 0.05 mM Cu^{2+} over the D39 strain + 0 mM Cu^{2+} was calculated from the measurements of at least 7 spots by Cyber-T.

For transcriptome analysis of D39 wild-type strain and its isogenic *copY* mutant, cells were grown in CDM without Cu^{2+} and harvested at an optical density at 595 nm of approximately 0.3. The experiments were repeated with four biological replicates. All other procedures regarding microarray were done as described above. Microarray data have been deposited to the Gene Expression Omnibus (GEO), and can be accessed via GSE30415.

Inductively coupled plasma- mass spectrometry (ICP-MS) analysis

For ICP-MS analysis, samples were prepared as describe before (224). In short, cultures of D39 wild-type and *copA* mutant were grown in 100 ml of CDM with and without 0.05 mM Cu^{2+}. Cultures were centrifuged and washed (at 4 °C) once with the CDM medium and twice with phosphate-buffered saline (PBS) that had been treated with chelex (Sigma) overnight. The cell pellets were dried overnight in a Speedvac. The dried cells were subsequently used for analysis by means of ICP-MS, as described before (105). Results were expressed as µg of Cu^{2+} g-1 dry weight of cells.

RNA extraction, reverse transcription (RT)-PCR and purification for quantitative RT-PCR

Total RNA was extracted from *S. pneumoniae* strain D39 grown to mid-log phase in Sicard's defined medium (229) in the presence or absence of 0.05 mM Cu^{2+} in microaerophilic conditions. Bacteria were harvested by centrifugation at 3000 g for 10 min. RNA was extractad by TRIZOL method as described by the manufacturer (Invitrogen, Paisley, UK), and purified by RNeasy purification kit (Qiagen, Crawley, UK). RNA was quantified and its integrity was checked by ethidium bromide staining after electrophoresis through a 1% (wt/vol) agarose gel. Any contaminating DNA was removed by treatment with 2 U RNase-free DNase I (Invitrogen, Paisley, United Kingdom) for 15 min at room temperature, followed by heat inactivation for 10 min at 65 °C in the presence of 2.5 mM EDTA.

First strand cDNA synthesis was performed on approximately 1 µg DNase-treated total RNA, immediately after isolation, using 200 U of SuperScript II reverse transcriptase (Invitrogen) and random hexamers at 42°C for 55 min. (272). cDNA (2 µl) was amplified in a 20 µl reaction volume that contained 1 x SYBR Green PCR master mix (Applied Biosystems, Foster City, USA) and 3 pmol of each primer (Table 2). The transcription level of specific genes was normalised to *gyrB* transcription, amplified in parallel with SP0806F and SP0806R primers. The reactions were performed in triplicate using the following cycling parameters: 1 cycle of 10 min 95 ºC followed by 40 cycles of 30 sec 95 ºC, 1 min 55 °C, and 30 sec 72 °C. The results were interpreted using the comparative C_T method (219). Differences in expression of two-fold or greater relative to control were considered as significant. To confirm the polycistronic nature of the *cop* operon total RNA was isolated from *S. pneumoniae* D39 wild-type grown in CDM + 0.05 mM Cu^{2+}. Primers copY-F and cupA-R were used to amplify the IR-I intergenic region between *copY* and *cupA,* whereas the IR-II intergenic region between *cupA* and *copA* was amplified using primers cupA-F and copA-R. Primers are listed in Table 2. PCRs were performed with 1/100 part of the RT reactions, and 200 ng of RNA and 45 ng DNA.

Extraction of pneumococcal RNA from infected tissues

Outbred 8 to 9-week-old female MF1 mice (Harlan Olac, Bicester, UK) were intranasally infected with 50 µl PBS containing $1X10^6$ passaged type 2 pneumococcal, as before (271). When the mice became severely lethargic they were anesthetised and blood

(0.5-1 ml) was collected by cardiac puncture. After killing by cervical dislocation, the lungs were removed and homogenized on ice in 10 ml sterile PBS using a tissue homogenizer. The pneumococcal mRNA was extracted from approximately 250-300 mg infected mouse lung tissue samples. The nasopharynx (40-60 mg/mouse) was dissected by removing the entire palate. Then nasopharyngeal tissues were transferred into sterile PBS and homogenised. To separate pneumococci from host cells, lung homogenates and blood samples were centrifuged at 900 g for 6 min at 4ºC. Supernatants were subsequently centrifuged at 15,500 g for 2 min at 4ºC, and the bacterial pellet was stored at -80ºC until further processing. Prior to pelleting, 20 µl homogenate was removed, serially diluted in PBS and plated onto blood agar in order to enumerate pneumococci and to exclude the presence of contaminating microorganisms. RNA extraction and purification were done as described in the previous section, and routinely 0.8-3 µg total RNA per sample could be obtained.

***In vivo* virulence studies**

Ten weeks old female MFI outbred mice (Harlan Olac), were used for virulence testing. A standardised inoculum was prepared as described previously (271, 273). To determine the virulence of pneumococcal strains, mice (n=10 for each group) were infected intranasally with approximately $1x10^6$ *S. pneumoniae* CFU as described before (271, 273). The inoculum dose was confirmed by viable counting on blood agar plates. Mice were monitored for disease signs (progressively starry coat, hunched, and lethargic) for 7 days, and those that reached the severely lethargic stage were considered to have reached the end point of the assay and were killed humanely. The time to this point was defined as 'survival time'. Mice that were alive 7 days after infection were deemed to have survived the infection. To determine the development of bacteraemia in each mouse, approximately 20 µl venous blood was obtained from intranasally infected mice at predetermined time points after infection. Viable counts in blood were determined by serial dilution in sterile PBS and plating onto blood agar plates (Oxoid, Basingstoke, UK) supplemented with 5% (v/v) defibrinated horse blood with appropriate antibiotic. Survival times were analysed by the Mann-Whitney U test. Growth of pneumococci in the nasopharynx and lungs was also determined at 0.5, 12, 24 and 36 h post infection. For this, at predetermined time intervals following intranasal infection, set groups of mice (n=5 for each time point) were deeply anesthetized as before and subsequently the mice were killed by cervical dislocation. The lungs and nasopharynx were transferred separately into 10 ml of sterile PBS, weighed, and homogenized (271). Viable counts in

homogenates were determined as above. Data were analysed by analysis of variance followed by the Beferroni post-test.

Results

S. pneumoniae encodes a number of putative copper homeostasis genes

Bioinformatic analysis showed that all sequenced pneumococcal genomes contain an operon with homology to the *cop*-like operon of *Enterococcus* and other species of the *Lactobacillale* order (208). The *S. pneumoniae cop*-like operon encodes a putative copper-dependent repressor protein (designated *copY*) and ATPase (designated *copA*) which are conserved in all other *Lactobacillale,* as well as a gene of unknown function which is unique to *S. pneumoniae*, *Streptococcus mitis* and *Lactobacillus johnsonii* (208) (Fig. 1A). Analysis of the flanking regions of the *cop*-like operon identified -10 and -35 promoter sequences in the upstream region of *copY* and a possible terminator sequence downstream of *copA* (Fig. 1A). RT-PCR using intergenic primer sets confirmed that the three genes form an operon and are transcribed as a single transcript, as is the case in *Enterococcus* (Fig. 1B) (233).

The *S. pneumoniae* CopY (SPD0633) polypeptide (Fig. 1A) has 33% identity and 64% similarity to *Enterococcus hirae* CopY at the amino acid sequence level. The N-terminal domain has all the conserved amino acids of the "winged" helix DNA binding motif found in other CopY proteins, but appears to have a truncated C-terminus and only has the first CxC part of the $CxCx_{4-6}CxC$ copper-binding domain found in other CopY homologues (197). The annotated translational start sites for *copY* differ between genome sequences. However analysis of the nucleotide sequences shows that there is 100% homology in this region suggesting that the genomes with *copY* translational start sites different to the D39 sequence shown in Fig. 1A are annotated incorrectly. *In silico* analysis identified two *lactobacillale* CopY binding motifs (TACAnnTGTA) in the *S. pneumoniae* D39 *copY* promoter region (Fig. 1A) (208), suggesting that the *S. pneumoniae cop* operon is auto-regulated by CopY.

The *copA* gene (*SPD0635*) encodes a P1-type ATPase, which has 35-44% amino acid sequence identity to other Gram-positive ATPases involved in the efflux of copper (208). The predicted CopA polypeptide contains three conserved domains; an amino terminal plastocyanin-like domain, an E1-E2 ATPase domain and a carboxyl terminal haloacid dehalogenase-like hydrolase domain. Unlike *copY* and *copA*, the third gene (*SPD0634*) in the *S. pneumoniae cop*-like operon is not highly conserved between species (208). *SPD0634* is

predicted to encode a 123-amino acid protein of unknown function, which has a plastocyanin or cupredoxin-like domain which may bind copper, and hence, has been designated *cupA*.

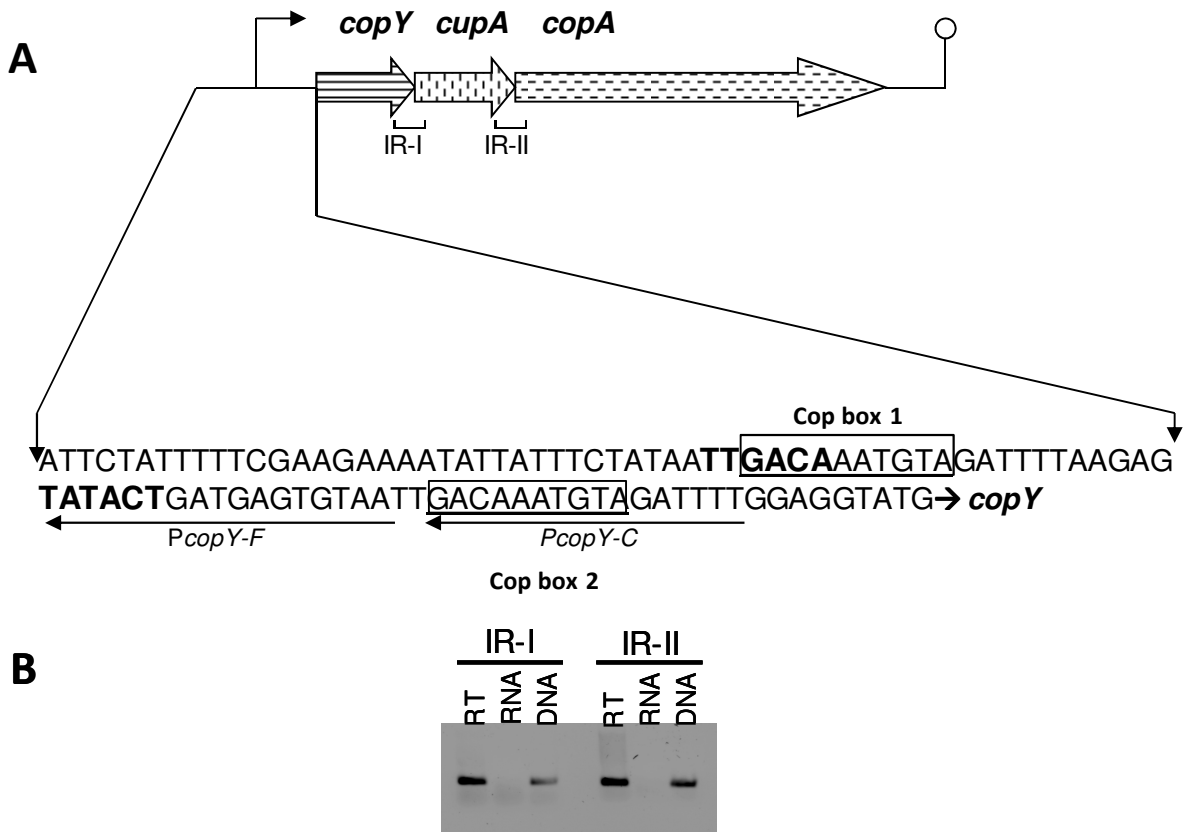


Figure 1: (A) The structure of the *S. pneumoniae cop* operon encoding *copY,* a putative copper transcriptional regulator, *cupA,* a hypothetical protein and *copA*, a copper-translocating P-type ATPase. The location of the putative promoter and terminator are indicated by an arrow and circle, respectively. Nucleotides in bold indicate the putative core promoter sequences and boxed nucleotides indicate putative *cop* box consensus sequences. Arrows indicate the sequence of PcopY-C and PcopY-F reverse primers used to construct P*copY-wt-lacZ* and P*copY-mut-lacZ.* **(B)** Reverse transcriptase PCR analysis to confirm the polycistronic nature of the *S. pneumoniae cop* operon. RT- PCR was performed on total RNA isolated from D39 wild-type grown in CDM + 0.05 mM Cu^{2+} with (RT) and without (RNA) reverse transcriptase treatment using the IR-I and IR-II intergenic region primer pairs (see Fig. 1A). DNA was used as a positive control.

Interestingly, unlike many other Gram-positives (75, 230, 233), *S. pneumoniae* does not encode a *copZ* copper chaperone. In Gram-positives, copper is usually donated to CopY by CopZ (50, 51). Therefore, it is not clear how copper is chaperoned or how the putative copper-dependent regulator CopY senses copper in pneumococci. *In silico* analysis only identified three other genes with homology to copper homeostasis proteins in the sequenced *S. pneumoniae* genomes - two P-type ATPase genes: *ctpE* (*SPD1927*) and *ctpC* (*SPD1436*); and gene *cutC* (*SPD1118*) which exhibits 34% amino acid identity and 53% similarity to the *E.*

coli cutC gene. CutC is required for maximal copper tolerance in *E. coli* (87) and is conserved in the *Lactobacillale* suggesting that it has an important function in these bacteria (208).

The *copY* promoter is induced by copper, repressed by zinc and auto-repressed by CopY

The transcriptional response of the *cop* operon to selected metal ions was investigated using *S. pneumoniae* D39 strain carrying an ectopic P*copY-lacZ* fusion integrated into the *bgaA* gene, which contained the entire *copY* promoter region. Specific β-galactosidase activity from the putative promoter was examined in CDM containing different concentrations of various metal ions: Zn^{2+}, Co^{2+}, Ni^{2+}, Fe^{2+} and Cu^{2+}. As seen in Fig. 2A, only Cu^{2+} ions caused induction of P*copY-lacZ* expression in *S. pneumoniae*. Further analysis of Cu^{2+} responsiveness of the promoter showed that there was a gradual response with increasing concentrations of added Cu^{2+} ions (Fig. 2B), with the highest level of expression observed at 0.05 mM Cu^{2+}, which is a concentration of Cu^{2+} that does not inhibit growth (Fig. 4).

The metal ion specificity of the *copY* promoter was further investigated by growing pneumococci with a constant amount of copper and various concentrations of other metal ions: Zn^{2+}, Co^{2+}, Ni^{2+} and Fe^{2+} (Fig. 3A). Addition of up to 0.05 mM Co^{2+} or Ni^{2+} had no effect on promoter activity. However, 0.1 and 0.5 mM Fe^{2+} caused a 1.4-fold increase ($P< 0.01$) and addition of 0.1 mM Zn^{2+} resulted in a 2-fold decrease ($P<0.01$) in promoter activity compared to that in the absence of any metal ions. The addition of increasing concentrations of zinc to various constant amounts of copper resulted in a concentration-dependent decrease in the expression of P*copY-lacZ* (Fig. 3B), suggesting that, as in *E. hirae* (50), zinc may act as a CopY co-repressor in *S. pneumoniae*.

To elucidate whether the *cop* operon is regulated by CopY, a *copY-stop* mutant derivative of D39 was constructed. Since *copY* is the first gene of the operon, a non-polar knock out was obtained through insertion of four premature stop codons in the *copY* coding sequences, in order to disrupt CopY translation, but not transcription of the *cop* operon. Transcriptional analysis showed that expression of P*copY-lacZ* was completely de-repressed in the *copY* mutant compared to the wild-type strain, demonstrating that the *cop* operon is auto-repressed by CopY (Fig. 3B and C). Increased expression of P*copY-lacZ* in the *copY* mutant compared to the wild-type strain in the presence of Cu^{2+}, demonstrates that there is a substantial level of CopY repression of the *cop* operon in the wild-type, even in the presence of Cu^{2+}. In the *copY*-stop mutant, the addition of zinc did not repress P*copY-lacZ* activity as seen in wild-type pneumococci (Fig. 3C). Complementation of the *copY* gene in the *copY*-stop mutant strain led to restoration of the repression and Cu^{2+}-responsiveness of P*copY* in the

copY-stop mutant (Fig. 3C). Overall, these results show that the *S. pneumoniae cop* operon is specifically induced by copper and is auto-regulated by the CopY repressor protein.

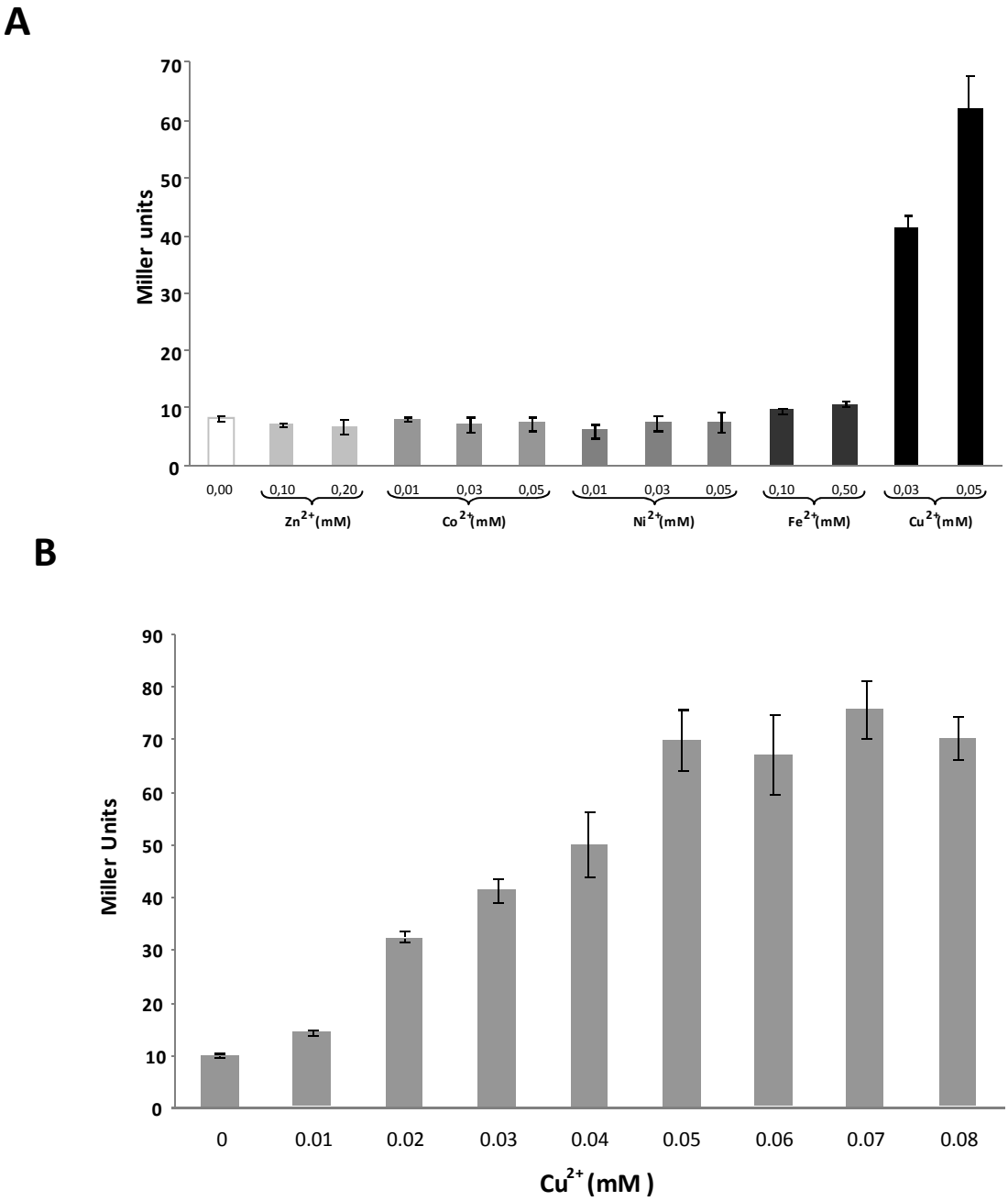


Figure 2: Expression levels (in Miller units) of a P*copY-lacZ* transcriptional fusion in D39 wild-type in CDM (without copper addition): **(A)** supplemented with various divalent metal ions; **(B)** supplemented with different concentrations of Cu^{2+}. Standard deviations of three replicates (A) or independent experiments (B) are indicated on each bar.

Two sequences similar to the *Lactobacillale cop* box consensus sequence are found in the *copY* promoter region of *S. pneumoniae* D39 (Fig. 1A). *Cop* box 1 is distal to the

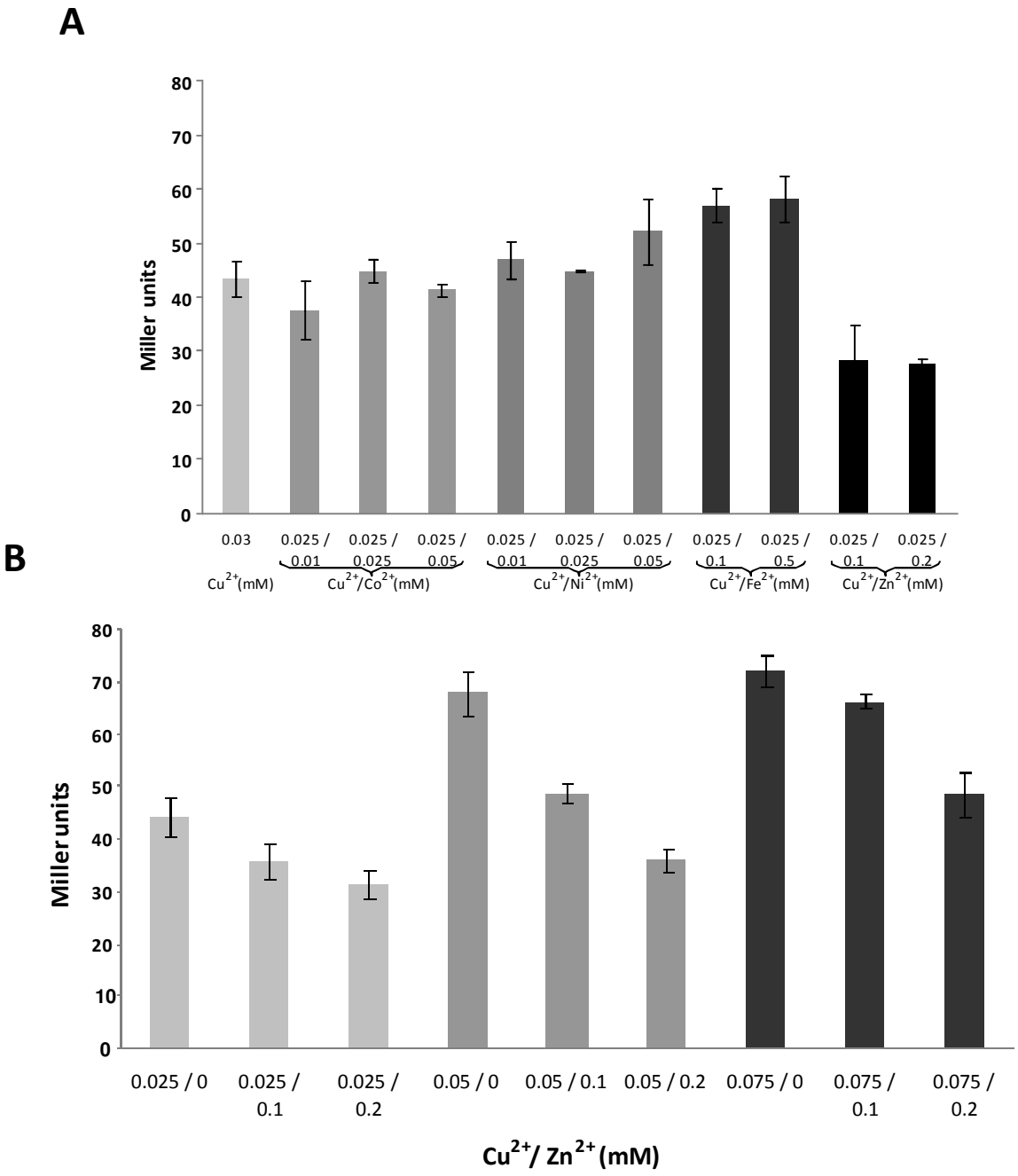
translational start sequence and overlaps with the predicted core promoter, while *cop* box 2 is present just upstream of the putative ribosomal binding site (RBS). The role of the CopY binding motif in copper-responsive gene expression was investigated using *S. pneumoniae* D39 strains carrying P*copY-lacZ* fusions that contained both *cop* boxes (P*copY-wt*) or with *cop* box 2 deleted (P*copY-mut*), which was expected not to interfere with the integrity of the core promoter. Specific β-galactosidase activity from the two promoters was compared in CDM supplemented with 0.05 mM Cu^{2+}. As seen in Fig. 3D, the absence of *cop* box 2 in P*copY-mut* resulted in a significant increase in β-galactosidase activity to the level observed with the wild-type promoter (P*copY-wt*) in the *copY-stop* mutant strain. In the *copY-stop* mutant, there was no significant difference observed between the wild-type and mutant promoters. Thus, these results show that CopY repression occurs through the *cop* box sequence identified in the *cop* operon promoter.

CopA is required for pneumococcal copper resistance *in vitro*

To investigate the physiological function of the *cop* operon genes in *S. pneumoniae*, isogenic *copY-stop, cupA-stop*, *ΔcopA* and *ΔcupA copA* mutants were constructed and characterised *in vitro*. Growth assays of wild-type D39 and the mutants in GM17 supplemented with various metal ions demonstrated that none of the strains tested appear to require Cu^{2+} for growth in these conditions (Fig. 4A). However, copper is toxic to pneumococci in high concentration as 1 mM Cu^{2+} was inhibitory to growth (Fig. 4A). In contrast, both the single *ΔcopA* and the double *ΔcupA copA* mutants showed significant inhibition of growth at 0.1 mM Cu^{2+} ($p<0.0$, Fig. 4A) and therefore demonstrated significantly decreased resistance against Cu^{2+} compared with the wild-type D39 and the *copY-stop* mutant. The *cupA-stop* mutant was also more sensitive to Cu^{2+} than the wild-type ($p<0.01$), although not to the same level as the *ΔcopA* single and double mutant suggesting that whilst the CupA copper binding protein plays a role in pneumococcal copper tolerance, the CopA ATPase is the major copper resistance mechanism.

The comparative growth of the wild-type D39 and *copY-stop* mutant and the difference in growth between the *cupA-stop* and *copA* mutants, show that the translational stop mutations in the *copY* and *cupA* genes do not prevent the expression of the downstream genes. The comparative growth of the *copY-stop* mutant and wild-type D39 is expected as de-repression of the copper efflux systems in the *copY-stop* mutant would be comparable to the induced response in the wild-type. No significant differences in growth were observed for any of the strains when grown with various concentrations of Co^{2+} and Ni^{2+} (Fig. 4C and D).

However wild-type D39 appears to grow significantly better than all the *cop* operon isogenic mutants in 0.1 and 0.2 mM Zn^{2+} ($p<0.01$) indicating that the *cop* operon may play a role in resistance to excess levels of Zn^{2+} as well as Cu^{2+}. There was no significant difference in the doubling times of the strains in the absence of added metal ions or in the presence of low copper/other metal ions (Zn, Ni and Co) as compared to the wild-type except for the *cupA, copA* and *cupA-copA* mutants which were increased in the presence of an increasing copper concentration (Fig. 4). 8 hour endpoints were taken as the important differences between all strains were most clearly visible after 8 hours of growth and no autolysis was observed at this time point. Therefore, together, our data show that CopA is specifically required for resistance to copper.



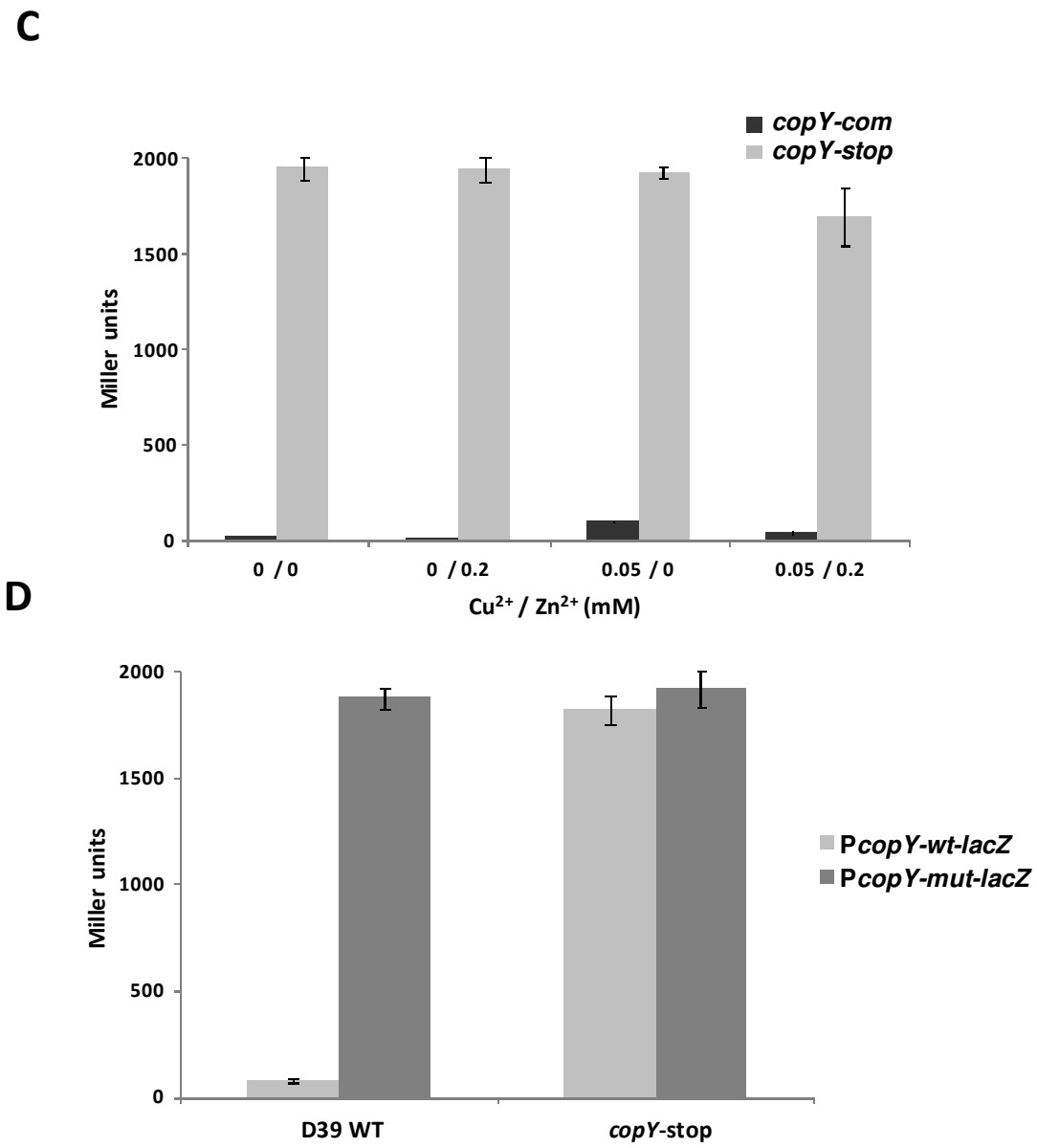


Figure 3: Expression levels (in Miller units) of a P*copY-lacZ* transcriptional fusion in D39 wild-type in CDM (without copper and zinc addition): **(A)** supplemented with various combinations of copper and zinc concentrations; **(B)** supplemented with 0.025 mM Cu^{2+} and with different concentrations of other divalent metal ions. Standard deviation of three independent experiments or replicates is indicated on each bar. **(C)** Expression levels (in Miller units) of a P*copY-lacZ* transcriptional fusion in D39 *copY*-stop and D39 *copY*-stop-com in CDM (without copper and zinc addition) supplemented with various combinations of copper and zinc concentrations. Standard deviation of three independent experiments is indicated on each bar. **(D)** Expression levels (in Miller units) of a P*copY-wt-lacZ* and P*copY-mut-lacZ* transcriptional fusions in D39 wild-type and D39 *copY*-stop in CDM supplemented with 0.05 mM Cu^{2+}. Standard deviation of three independent experiments is indicated on each bar.

ICP-MS analysis with and without added 0.05 mM Cu^{2+}

To determine whether the transcriptional effects observed in high and low copper microarray, correlate with a cell-associated concentration of Cu^{2+}, ICP-MS analysis was

performed on cells grown in CDM medium without and with 0.05 mM added Cu^{2+}. ICP-MS analysis revealed that *S. pneumoniae* D39 grown in the presence of added Cu^{2+} has a 13 fold higher cell-associated amount of Cu^{2+} as compared to D39 grown in the absence of added Cu^{2+} (13 μg g-1 dry mass of cells versus >1 μg g-1 dry mass of cells). These results indicate that the transcriptome effects observed above are mainly due to the difference in intercellular Cu^{2+} concentration.

To determine the role of CopA, the *copA* mutant was grown in CDM medium with 0.05 mM added Cu^{2+}. ICP-MS analysis revealed high accumulation of intracellular Cu^{2+} in the CopA mutant as compared to the D39 wild-type in the presence of 0.05 mM added Cu^{2+} (65 μg g-1 dry mass of cells versus 13 μg g-1 dry mass of cells. High accumulation of Cu^{2+} in the *copA* mutant suggests the function of CopA as a Cu^{2+} efflux transporter in *S. pneumoniae*.

Identification of copper regulated genes in *S. pneumoniae*

To investigate the effect of copper on global gene expression in *S. pneumoniae* and the role that CopY plays in that response, the transcriptomes of (i) D39 wild-type grown in CDM with and without copper, and (ii) D39 wild-type and D39 *copY* mutant strains in low copper were compared. For (i), a concentration of 0.05 mM Cu^{2+} was used as this concentration does not significantly affect growth but activates P*copY* to a high degree (Figs. 2 and 5). The amount of zinc in CDM was decreased to 5 μM from the standard concentration of 17.5 μM, in order to prevent Zn^{2+} repression of copper responsive genes as was seen in Fig. 3.

Tables 3 and 4 summarize the transcriptomic changes in *S. pneumoniae* induced by copper excess compared to limitation, in D39 wild-type. After applying the criteria of $\geq$ 2.0-fold difference as the threshold change and a *P* value < 0.001, 39 genes were differentially expressed, of which 22 were up- and 17 were down-regulated in the presence of copper. The *copY* (SPD0633)*, cupA* (SPD0634) and *copA* (SPD0635) genes showed the highest level of expression in high copper conditions, confirming the β-galactosidase assays (Figs 2 and 3) and also demonstrating that the microarray results are a valid representation of the pneumococcal global response to excess copper (Table 3). Strikingly, the transcriptional profiling did not identify any of the known copper homeostasis homologues mentioned above (*ctpC, ctpE* and *cutC*). However, several genes displayed increased expression in the presence of copper, including: a number of putative exoglycosidases such as the surface-associated β-N-acetyl glucosaminidase (SPD0444) and *strH* which is an important pneumococcal virulence factor (122)a putative operon encoding an uncharacterized transcriptional regulator (SPD1565) and a putative thioredoxin, and several amino acid transporters (Table 3). Unlike

the effect of copper in other Gram-positive bacteria, such as *M. tuberculosis* (261) and *Staphylococcus aureus* (18), there was no induction of any genes known to be involved in oxidative stress resistance or the misfolded protein response, showing that *S. pneumoniae* responds differently to copper than other bacteria.

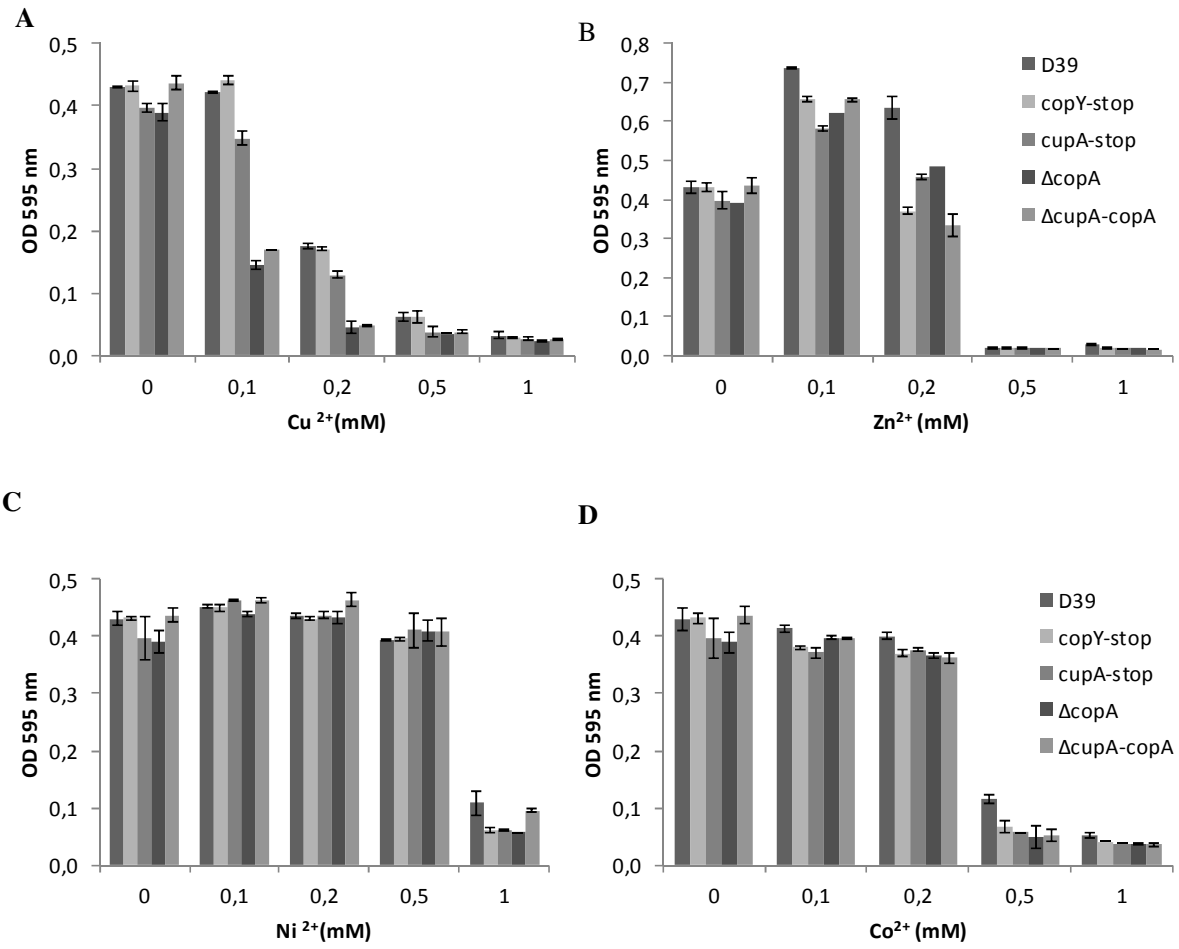


Figure 4: Growth of *S. pneumoniae* strains D39 and the isogenic *copY-stop*, *cupA*-stop, *ΔcopA* and *cupA-copA*$^-$ mutants after 8 hrs of growth in GM17 with increasing concentrations of Cu^{2+} **(A)**, Zn^{2+} **(B)** Ni^{2+} **(C)** or Co^{2+} **(D)**. Optical densities at 595 nm were determined after 8h of growth. Results represent the mean and standard deviation of three independent experiments.

Genes down-regulated in the presence of copper included several constituents of the purine metabolism pathway, SPD0051, SPD0058, SPD0059, SPD1628 and SPD1629, an operon encoding proteins involved in glycerophospholipid metabolism, as well as a number of ABC transporters (Table 4). Several hypothetical proteins, without predicted functions, were also among the differentially expressed genes.

Table 3: Summary of up-regulated genes in transcriptome comparison of *S. pneumoniae* strain D39 grown in CDM plus 0.05 mM Cu^{2+} and CDM plus 0 mM Cu^{2+}.

TIGR4 locus tag	D39 locus tag	Function (TIGR4 annotation)	Ratio[a]
SP0057	SPD0063	N-acetyl-β-hexosaminidase	2.7
SP0090	SPD0088	ABC transporter, permease protein	4.4
SP0148	SPD0150	ABC transporter, substrate-binding protein	2.1
SP0409	SPD0373	Hypothetical protein	3.3
SP0498	SPD0444	Endo-β -N-acetylglucosaminidase	3.1
SP0620	SPD0540	Amino acid ABC transporter, amino acid-binding protein	2.1
SP0709	SPD0616	(polar) Amino acid ABC transporter, ATP-binding protein	2.1
SP0710	SPD0617	(polar) Amino acid ABC transporter, permease protein	2.7
SP0711	SPD0618	(polar) Amino acid ABC transporter, permease protein	3.3
SP0727	SPD0633	Putative copper responsive regulator	19.0
SP0728	SPD0634	Hypothetical protein	19.1
SP0729	SPD0635	Putative copper-transporting P-type ATPase	11.6
SP1774	SPD1565	Transcriptional regulator (ArsR family)	1.5
SP1775	SPD1566	Hypothetical protein	2.3
SP1776	SPD1567	Thioredoxin	2.0
SP2072	SPD1899	Glutamine amidotransferase, class-I	2.7
SP2132	SPD1962	Hypothetical protein	3.7
SP2133	*	Hypothetical protein	2.8
SP2141	SPD1969	Glycosyl hydrolase-related protein	3.3
SP2142	SPD1970	ROK family protein	4.4
SP2143	SPD1971	alpha-mannosidase	3.2
SP2144	SPD1972	hypothetical protein	4.2

[a]Ratios ≥2.0 (D39 wild-type + 0.05 mM Cu^{2+} compared to D39 wild-type + 0 mM Cu^{2+}). P values all <0.001. In case of putative operons neighbouring genes with ratios <2.0 are also indicated.
* not annotated in NCBI database, despite 100% DNA sequence identity to TIGR4 SP2133
| putative operons

The transcriptome analysis of the *copY*-stop mutant strain showed strong up-regulation of the *cop* operon compared to the wild-type strain, thus confirming the auto-regulation of this transcriptional unit (Table 5). Other genes affected by the *copY-stop* mutation include genes participating in cellobiose utilisation (SPD0277, SPD0280 and SPD0282) which were up-regulated in the *copY*-stop mutant. Four genes were down-regulated (SPD1284, SPD1263-5), two of which encode a putative ABC transporter. Surprisingly, very few of the copper-responsive genes described in Tables 3 and 4 appeared to be affected by the *copY* mutation in *S. pneumoniae,* suggesting that CopY may not be the only copper-responsive regulator in *S. pneumoniae* or that these genes may be stimulated indirectly by the presence of copper. Therefore, it remains to be investigated how most of the copper-responsive genes identified by the microarray are regulated.

Table 4: Summary of down-regulated genes in transcriptome comparison of *S. pneumoniae* strain D39 grown in CDM plus 0.05 mM Cu^{2+} and CDM plus 0 mM Cu^{2+}.

TIGR4 locus tag	D39 locus tag	Function (TIGR4 annotation)	Ratio[a]
SP0044	SPD0051	Phosphoribosylaminoimidazole-succinocarboxamide synthase	-34.1
SP0051	SPD0058	Phosphoribosylamine-glycine ligase	-4.7
SP0053	SPD0059	Phosphoribosylaminoimidazole carboxylase catalytic subunit	-2.3
| SP0112	SPD0109	Amino acid ABC transporter, periplasmic amino acid-binding protein	-2.2
| SP0113	SPD0110	Argininosuccinate synthase	-2.7
SP0287	SPD0267	Xanthine/uracil permease family protein	-2.6
SP0585	SPD0510	Homocysteine methyltransferase	-3.2
SP1027	SPD0913	Inosine-5'-monophosphate dehydrogenase	-2.1
SP1127	SPD1010	Hypothetical protein	-2.2
| SP1695	SPD1506	Acetyl xylan esterase	-2.3
| SP1696	*	Hypothetical protein	-5.4
SP1826	SPD1609	ABC transporter, substrate-binding protein	-2.0
| SP1847	SPD1628	Xanthine phosphoribosyltransferase	-4.7
| SP1848	SPD1629	Xanthine permease	-2.8
SP2184	SPD2011	Glycerol uptake facilitator protein	-2.1
SP2185	SPD2012	Hypothetical protein	-2.0
| SP2186	SPD2013	Glycerol kinase, GlpK	-2.4
SP2240	SPD2069	Sporulation protein, SpoJ	-2.2

[a]Ratios ≤-2.0 (D39 wild-type + 0.05 mM Cu^{2+} compared to D39 wild-type + 0 mM Cu^{2+}). P values all <0.001
In case of putative operons neighbouring genes with ratios >-2.0 are also indicated.
* not annotated in NCBI database, despite 100% DNA sequence identity to TIGR4 genome
| putative operons

Table 5: Genes differently expressed in D39 *copY*-stop strain grown in CDM with no added Cu^{2+}.

TIGR4/R6 locus tag	D39 locus tag	Function (TIGR4 annotation)	Ratio[a]
SP0303	SPD0277	6-phospho-β-glucosidase, CelA	2.6
SP0306	SPD0280	Putative transcriptional regulator, CelR	2.4
SP0308	SPD0281	PTS system, IIA component, CelC	1.5
| SP0309	SPD0282	Hypothetical protein	2.7
| SP0727	SPD0633	Putative copper responsive regulator	59.7
| SP0728	SPD0634	Hypothetical protein	42.9
| SP0729	SPD0635	Putative copper-transporting P-type ATPase	30.0
SPR1309	SPD1284	Hypothetical protein	-3.0
\| SP1434 | SPD1263 | ABC transporter, ATP-binding/permease protein | -2.0 |
\| SP1435 | SPD1264 | ABC transporter, ATP-binding/permease protein | -1.6 |
\| SP1436 | SPD1265 | Hypothetical protein | -1.3 |

[a]Ratios ≥2 or ≤-2.0 (D39 *copY*-stop compared to D39 wild-type). P values all < 0.001.
In case of putative operons neighbouring genes with ratios <2.0 and >-2.0 are also indicated.
| putative operons

qRT-PCR was used to confirm data from the global transcript analysis. Fig. 5A confirms that expression of the *copY* and *copA* genes is induced in pneumococci grown in the presence of 0.05 mM Cu^{2+} and also demonstrates that, in agreement with the transcriptional

profiling, the putative *ctpC, ctpE* and *cutC* genes do not show copper-dependent expression *in vitro.* Thus, it appears that CopA is the main *S. pneumoniae* copper transporter induced under these growth conditions.

The *cop* operon is induced in pneumococci isolated from the lungs and nasopharynx but not the blood

We also investigated the expression of *copY*, *copA*, *ctpC, ctpE* and *cutC* genes in pneumococci recovered from infected mouse tissues. None of the five genes tested showed any difference in expression level in pneumococci isolated from the blood compared to pneumococci grown *in vitro* in Sicard's medium (Fig. 5B). However, all five genes demonstrated induced expression in the lungs (Fig. 5B), with *copA* and *cutC* showing higher levels of expression than the other genes. In addition, both *copA* and *cutC* expression showed an increase in pneumococci isolated from the nasopharynx compared to *in vitro*. These results also show that *copA* expression is significantly higher than *copY in vivo* even though the two genes are in a putative operon, suggesting that the regulation of these genes may be complex. Consequently, these results suggest that CopA and CutC may be important for pneumococcal growth in the nasopharynx, which has not been reported previously.

***S. pneumoniae* CopA is required for copper resistance *in vitro* and for growth in the nasopharynx and lungs in a mouse model of pneumonia**

To determine the role of CopA in copper tolerance *in vivo*, a *S. pneumoniae* D39 *copA*$^-$ insertion mutant was constructed and characterised. *In vitro* the *copA*$^-$ mutant showed a growth defect in excess copper concentrations similar to the D39 *ΔcopA* and *ΔcupA copA* strains shown in Fig. 4 (data not shown), therefore the two different *copA* mutant strains have the same *in vitro* phenotype.

In a mouse model of pneumonia following intranasal infection it was found that the *copA*$^-$ mutant was less virulent than the D39 parental strain. The median survival time of the *copA*$^-$ mutant infected group (58 ±7 h) was significantly longer than the wild-type infected group (47 ±4 h) ($p<0.01$) (Fig. 6A). While the wild-type could be detected in blood at 8 h after infection, the mutant was only detected 8 to 12 h post-infection (Fig. 6B). However, once in blood, the growth patterns of the strains were similar. The numbers of pneumococci also were monitored in the nasopharynx and lungs at different time points (Fig. 6C and D). In the nasopharynx, the *copA*$^-$ mutant colony counts were less than D39 at 12, 24, and 48 h post-infection ($p<0.01$ for 12h and $p<0.001$ for 24 h and 48 h). In the lungs, a difference in the

numbers of the *copA*$^-$ mutant and the wild-type was only detected at 24 h post-infection when the *copA*$^-$ mutant numbers (2.8 ±0.2) were significantly lower than D39 (4.0 ±0.1) ($p<0.05$). These results are consistent with the expression data, which showed that *copA* expression was higher in the lungs and nasopharynx. Copper growth assays of wild-type D39 and the *copA*$^-$ mutant strains isolated from the mice showed that reversion of the *copA*$^-$ mutation and accumulation of secondary mutations had not occurred during passage of the pneumococci *in vivo,* as the passaged bacteria had the same copper resistance phenotype as the bacteria prior to inoculation (data not shown). Therefore, together, these data demonstrate that copper homeostasis is important for pneumococcal survival in the lungs and nasopharynx.

Discussion

In this study, we showed that copper-responsive gene regulation and resistance mechanisms are important for pneumococcal physiology, growth in the nasopharynx and virulence during pneumonia.

Our data showed that the pneumococcal response to copper differs to most other bacteria. Bacteria usually initiate a global adaptive genetic response to copper, which involves induction of other stress regulons (18, 118, 238, 261). The pneumococcal copper regulon is surprisingly smaller than many other bacteria as only 39 *S. pneumoniae* genes are differentially expressed in the presence of copper compared to approximately 300 genes in other pathogens such as *Enterococcus faecalis* (208) and *Pseudomonas aeruginosa* (118, 238). In addition, unlike other Gram-positive bacteria, such as *Mycobacterium tuberculosis* (261) and *Staphylococcus aureus* (18), there is no induction of any genes known to be involved in oxidative stress resistance or the misfolded protein response, showing that *S. pneumoniae* responds differently to copper than other bacteria. This may be typical for lactic acid bacteria because in *Lactococcus lactis* only 14 genes were shown to be copper-regulated (157). Therefore, these differences in copper regulons may reflect the physiological requirements and environmental niches of the different bacteria.

Our data show that the *cop* operon is repressed by CopY in low copper conditions. This demonstrates that the pneumococcal CopY protein responds to copper even though it only has the first CxC part of the $CxCx_{4\text{-}6}CxC$ copper-binding domain found in other CopY homologues (197). The mechanism of CopY copper-responsive repression in *Enterococcus* has been well studied (197), and has been shown to be a global regulator (208) as has the *Lactococcus lactis* CopY homologue CopR (157). Therefore, it was surprising that in the *S.*

pneumoniae copY-stop mutant only the *cop* operon expression was significantly increased compared to the wild-type, suggesting that CopY is not a global regulator in *S. pneumoniae*. The other copper-responsive genes identified by our transcriptional profiling either may be directly regulated by copper via an unknown regulator or are part of other regulons that are stimulated indirectly by the presence of copper.

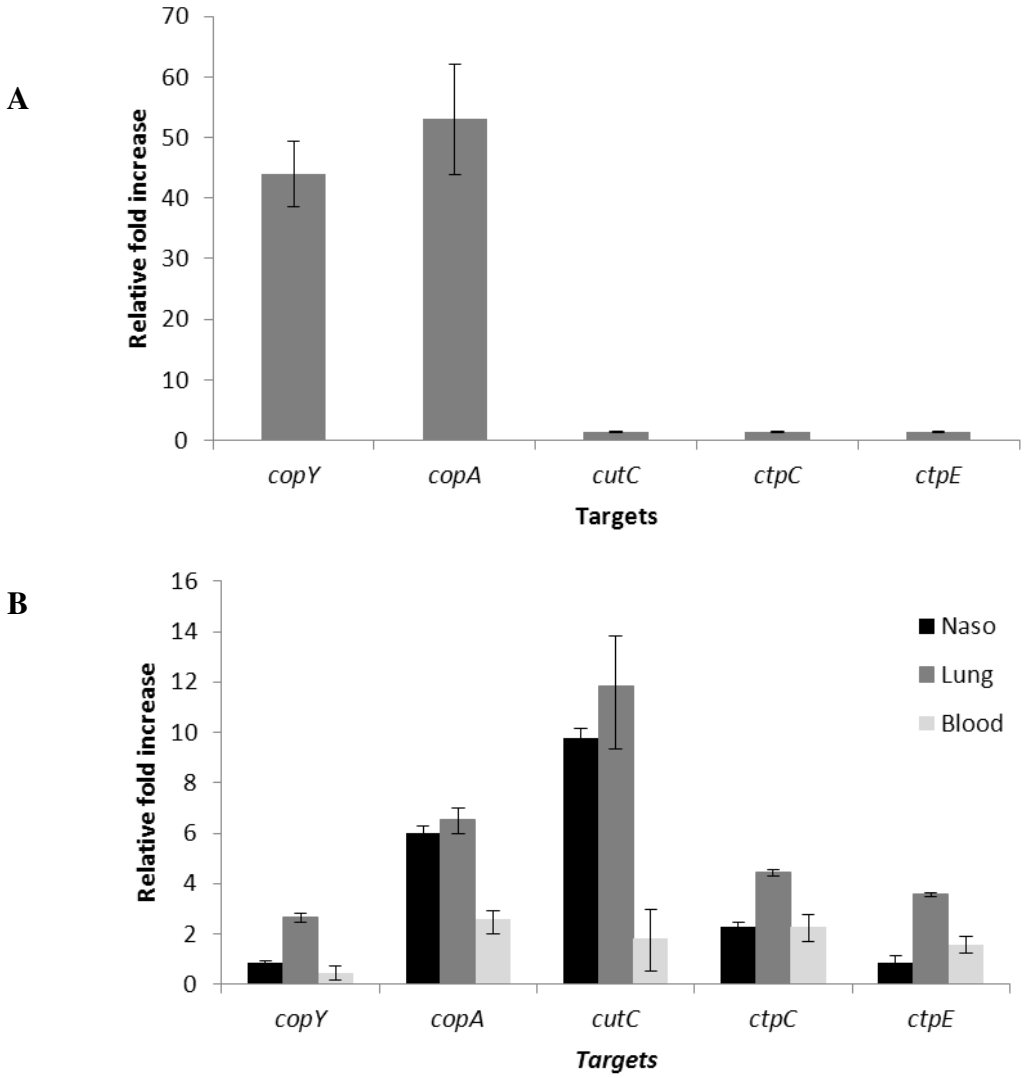


Figure 5: The relative *in vitro* and *in vivo* expression of pneumococcal genes linked to copper metabolism. **(A)**. *S. pneumoniae* strain D39 was grown in Sicard's defined medium *in vitro*, in the presence or absence of 0.05 mM Cu^{2+}. The expression of the genes *copY*, *copA*, *cutC*, *ctpC* and *ctpE* was normalised with housekeeping gene *gyrB.* Results represent the mean and standard deviation of three independent experiments. The fold increase is relative to the expression in Sicard's in the absence of copper. **(B)** Total RNA was extracted from D39 pneumococci isolated from the nasopharynx (black bars), lungs (dark grey bars), and blood (light grey bars) of intranasally infected MFI mice. The expression of the genes *copY*, *copA*, *cutC*, *ctpC* and *ctpE* was normalised with housekeeping gene *gyrB.* Results represent the mean and standard deviation of three independent experiments. The fold increase is relative to the expression in Sicard's in the absence of copper *in vitro.*

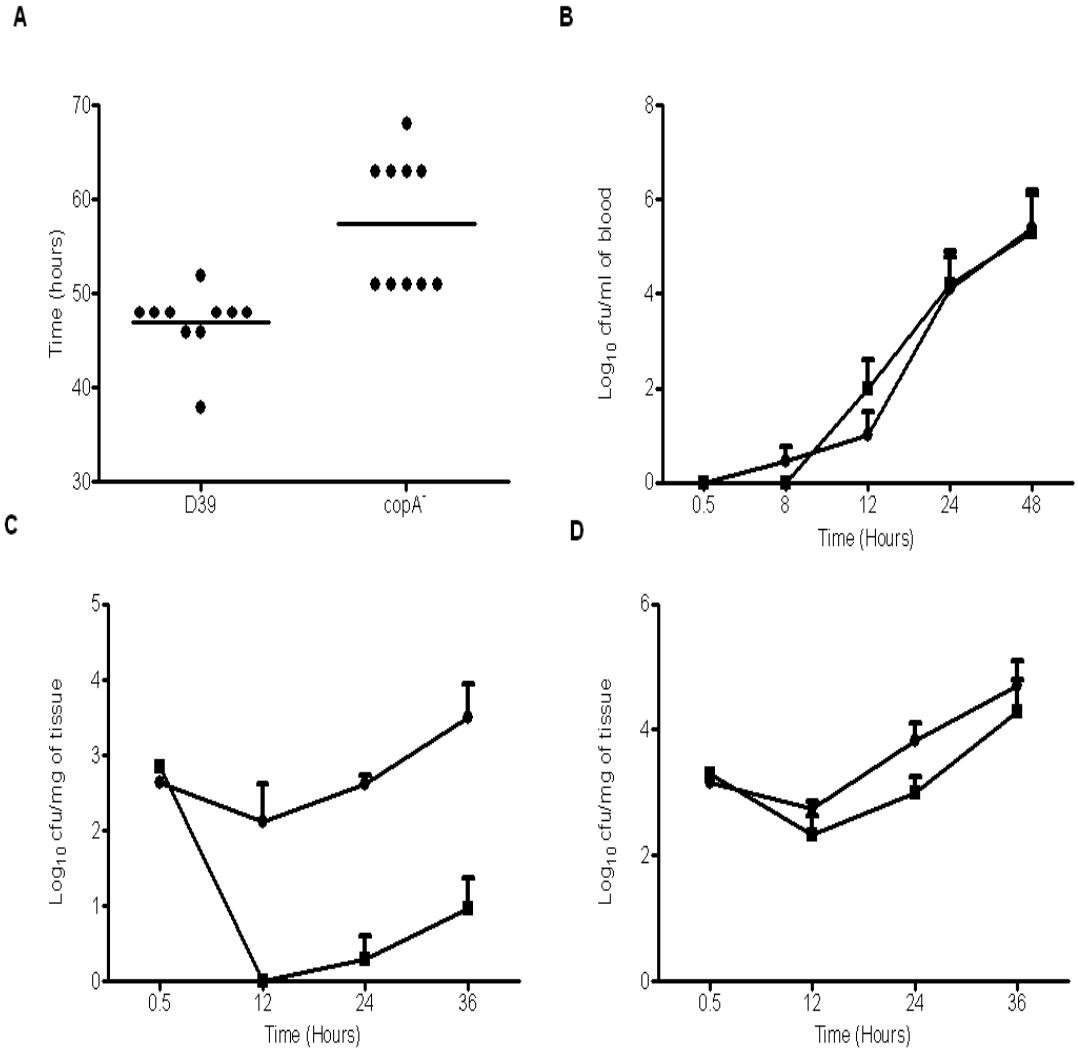


Figure 6. (A) Survival of mice infected intranasally with D39 or *cop mutant*. Symbols show the time that individual mice became severely lethargic. The horizontal bars mark the median time to the severely lethargic state. Growth of D39 (●) and *copA*$^-$ (■) in **(B)** blood, **(C)** nasopharynx and **(D)** lungs. Survival time after intranasal infection with D39 (●) and *copA*$^-$ (■). Each datum point derived from 10 mice. Vertical bars indicate SEM.

Bioinformatic analysis of the D39 wild-type genome with the *S. pneumoniae* P*copY cop* box sequence (GACAAATGTA, Fig. 1A) identified only two perfect matches in intergenic regions; the two *cop* boxes located in P*copY,* which agrees with the microarray data showing that CopY only regulates the *cop* operon. Searching the *S. pneumoniae* D39 genome sequence with the *cop* box sequence (GACAAATGTA) allowing 1 mis-match yielded 130 genes. However, none of these genes are regulated by CopY in the microarray analysis. Previous publications have shown that many of the genes identified by motif searches are not CopY-regulated when investigated experimentally (157). It may be that the consensus sequence used for the bioinformatic analysis is not the minimum sequence necessary for CopY regulation in *S. pneumoniae* and that more bases may be essential or these genes are not

regulated by CopY solely and so are still repressed by other factors under the growth conditions used for the microarray.

Copper is thought to become toxic to bacterial and eukaryotic cells through two major mechanisms: unliganded Cu^{2+} conversion to Cu^{1+} which reacts with H_2O_2 catalysing the generation of highly toxic hydroxyl radicals ($OH^{\bullet}$) and direct interaction of copper with cellular molecules (13). An important copper toxicity mechanism in *Escherichia coli* is the inactivation of the iron-sulphur clusters of the dehydratase enzymes which leads to defective branched chain amino acid biosynthesis (156). Therefore, most organisms have evolved a number of mechanisms to counteract and prevent further $OH^{\bullet}$ generation as well as limit free copper in the cell.

Copper toxicity is counteracted in *S. pneumoniae* through the action of the *cop* operon. Like other bacteria, our data shows that *S. pneumoniae* induces a response to protect the cell from free copper, which involves the CopA ATPase protein. The unique pneumococcal transcriptional response to copper suggests that copper toxicity in *S. pneumoniae* may also involve previously undefined mechanisms. *S. pneumoniae* is already well adapted to oxidative stress as it produces mM levels of H_2O_2 that are toxic to other bacteria (190), which may explain the lack of an induced oxidative stress response in the microarray upon exposure to copper. In addition, pneumococci are auxotrophic for several branched chain amino acids due to incomplete biosynthetic pathways (116) and have to rely on environmental sources of branched amino acids. This means that pneumococci may be more resistant to copper toxicity mediated through the inactivation of the iron-sulphur clusters of the dehydratase enzymes (156). Interestingly a number of pneumococcal amino acid transporters are up-regulated on exposure to copper, suggesting that there is a requirement for these amino acids in response to the presence of copper possibly due to the toxicity of copper for the biosynthetic pathways of these amino acids.

It is now becoming apparent from our data that copper homeostasis is important for the survival of *S. pneumoniae* in specific host sites. There is increasing evidence that copper is an important metal in the lungs. Copper concentrations are higher in the lungs compared to blood (46). Concurrent with this, our data show increased expression of *copY, copA, cutC, ctpE* and *ctpC* and the attenuation of virulence of a *copA*$^{-}$ mutant in the lungs but not in blood. In addition, signature tagged mutagenesis identified CopA as well as CtpE and CutC as being important for pneumococcal infection of the lung (93). Copper also appears to be important for the behaviour of pneumococcus in the nasopharynx as there is increased expression of *copA* and *cutC,* and decreased survival of the *copA*$^{-}$ mutant in the nasopharynx in this mouse

model of acute invasive disease. CutC is conserved in other *Lactobacillale* species suggesting that it may have an important role in copper homeostasis (206). The decrease in the *copA*$^-$ mutant bacterial count in the nasopharynx could be due to the bacteria being unable to adhere to the cells or mucus of the nasopharynx, or the *copA*$^-$ mutant bacteria may be unable to exploit the environment and grow in the nasopharynx. The importance of copper resistance for virulence has been shown for other respiratory pathogens including *Mycobacterium tuberculosis* (268) and *Pseudomonas aeurginosa* (220). However, the importance of bacterial copper homeostasis in the nasopharynx has not been described previously and requires further investigation.

Acknowledgements

This study was supported by the Wellcome Trust to JAM and PWA (076573/Z/05/Z) and to HY and PWA (078763/Z/05/Z), and the Higher Education Commission (HEC) Pakistan to Sulman Shafeeq. The research leading to these results has also received funding awarded to PWA from the European Community's Seventh Framework Programme FP7/2007-2013 under grant agreement number HEALTH-F3-2009-222983. We thank Anne de Jong and Siger Holsappel for help with the DNA microarray slide production.

Chapter 4

CelR-mediated activation of the cellobiose-utilization gene cluster in *Streptococcus pneumoniae*

Sulman Shafeeq, Tomas G. Kloosterman and Oscar P. Kuipers

Based on *Microbiology* **157**: 2854-61 (2011).

Abstract

The human pathogen *Streptococcus pneumoniae* harbors many genes encoding phosphotransferase systems (PTSs) and sugar ABC (ATP-binding cassette) transporters, including systems for the utilization of the β-glucoside sugar cellobiose. In this study, we show that the transcriptional regulator CelR, which has previously been found to be important for pneumococcal virulence, activates the expression of the cellobiose-utilization gene cluster (*cel* locus) of *S. pneumoniae*. Expression directed by the two promoters present in the *cel* locus was increased in the presence of cellobiose as a sole carbon source in the medium, while expression decreased in the presence of glucose in the medium. Furthermore, we have predicted a 22-bp putative CelR regulatory site (5'-YTTTCCWTAWCAWTWAGGAAAA-3') in the promoters of *celA* and *celB*, and *in silico* analysis showed that it is highly conserved in other pathogenic streptococci as well. Promoter truncations of *celA* and *celB*, where the half or full CelR regulatory site was deleted, confirmed that the CelR binding site in P*celA* and P*celB* is functional. Transcriptome studies with the *celR* mutant and *in silico* prediction of the CelR regulatory site in the entire D39 genome sequence show that the *cel* locus is the only cluster of genes under the direct control of CelR. Therefore, CelR is a regulator dedicated to the cellobiose-dependent transcriptional activation of the *cel* locus.

Introduction

The major Gram-positive human pathogen *Streptococcus pneumoniae* has the ability to colonize the nasopharynx (31) and from there spread to different parts of the human body causing serious infections like pneumonia, meningitis, otitis media and sepsis (114, 182). For *S. pneumoniae*, a big challenge for survival in the host environment is the adaptation to the changing nutritional conditions encountered in the different niches inside the host, according to which it has to tune the expression of the appropriate genes (39, 172, 228).

Glucose is the preferred source of energy for many bacteria, but the presence of different sugar-specific PTSs gives them the ability to utilize other carbon sources as well (68). Bacteria are able to ferment different β-glucosides such as cellobiose, aesculin, arbutin and salicin, which are commonly present in plants (119). Cellobiose can be a potential source of energy for different streptococci as well (118, 168, 275). In *S. pneumoniae*, the gene cluster for cellobiose utilization (*cel* locus) is comprised of seven genes encoding a 6-phospho-β-glucosidase (CelA), a putative DNA-binding transcriptional regulator (CelR), the A, B and C domains of a cellobiose PTS EII^{Cel} (CelB, CelC, and CelD) and two hypothetical proteins. It

has been shown that *S. pneumoniae* lacking the *celR, celD* or the hypothetical gene *spr0281* (*SPD0282*) gene was hardly able to grow in a chemical defined medium (CDM) supplemented with cellobiose as a sole carbon source (168). A similar study has been performed in *Streptococcus mutans* (275), where strains with deletion of *celA*, *celB, celC, celD* or *celR* were unable to grow in medium with cellobiose as a sole source of energy.

Interestingly, the extracellular matrix (ECM) of mammalian tissues is rich in glycosaminoglycans (GAGs) that contain repeating units of β-linked disaccharides (119). The degradation of GAGs from the mammalian ECM may release structural analogues of cellobiose (119). Although it is unlikely that *S. pneumoniae* encounters cellobiose in the human body, it is possible that it can utilize structurally similar β-glucosides that are derived by degradation of the mammalian ECM (119, 121).

The two-component system TC08 is one of the 13 TCSs (244) found in *S. pneumoniae* and has been shown to be involved in the regulation of the pilus locus in strain TIGR4 (234). In strain R6, it regulates the *cel* locus (168), as a mutant producing a hyperactive variant of HK08 displayed high repression of transcription of the *cel* genes (168). Moreover, this *hk08* mutant was not able to grow in the presence of cellobiose as a sole energy source (168). On the other hand, in *S. mutans* the *cel* locus was activated by the transcriptional regulator CelR and a strain lacking *celR* was unable to grow in medium with cellobiose as a sole energy source (275). Furthermore, the *celA* gene in the *cel* locus encodes one of the three β-glucosidases present in *S. mutans* and a strain in which *celA* was deleted was unable to ferment cellobiose, amygdalin, gentobiose and salicin (186). Deletion of the other two β-glucosidases in *S. mutans* had no effect on the fermentation of these four β-glucosides (186). These studies demonstrate the importance of CelA in the fermentation of cellobiose and the role of CelR as an activator of the *cel* locus in *S. mutans*.

In this study, we show that CelR acts as a cellobiose-dependent transcriptional activator of the *cel* locus in *S. pneumoniae* as well. Furthermore, in the presence of cellobiose or glucose, the expression of the *cel* locus was not affected by deletion of either *hk08* or *ccpA*. Investigation of the *celR* deletion strain by means of DNA microarray analyses demonstrated that, besides the *cel* locus, CelR has no other targets under the conditions tested. By MEME motif search, a putative regulatory site of CelR was identified in the promoter regions of *celA* and *celB*, which is conserved in P*celA* and P*celB* of other streptococci as well. The functionality of this CelR regulatory site was proven by means of promoter truncation, showing that deletion of the half or full regulatory sequence leads to total inactivation of the *celA* and *celB* promoters.

Material and Methods

Bacterial strains and growth conditions

Bacterial strains and plasmids used in this study are listed in Table 1. M17 broth (242) supplemented with 0.5% (w/v) glucose was used for growing *S. pneumonia*e D39 wild-type (136) on blood agar plates supplemented with 1% (v/v) defibrinated sheep blood in micro-aerophilic conditions at 37^{o}C. For β-galactosidase assays, derivatives of *S. pneumoniae* D39 were grown in M17 medium supplemented with different sugars (Glucose and Cellobiose with a concentration (w/v) as mentioned in the Results section, and cells were harvested at mid-exponential phase. For selection on antibiotics, media were supplemented with the following concentrations of antibiotics: erythromycin: 0.25 μg ml^{-1} for *S. pneumoniae* and 120 μg ml^{-1} for *E. coli*; spectinomycin: 150 μg ml^{-1} and tetracycline: 2.5 μg ml^{-1} for *S. pneumoniae*; and ampicillin: 100 μg ml^{-1} for *E. coli*. All bacterial strains used in this study were stored in 10% (v/v) glycerol at -80 °C.

DNA isolation and manipulation

All DNA manipulations in this study were done as described before (125). For PCR amplification, chromosomal DNA of *S. pneumoniae* D39 wild-type (136) was used. Primers used in this study are listed in Table 2 and based on the sequence of the D39 genome (136).

Construction of deletion mutants of *hk08* and *celR*

hk08 and *celR* deletion mutants were made by allelic replacement with an erythromycin- and spectinomycin-resistance marker, respectively. Briefly, primers HK08-1/HK08-2, HK08-3/HK08-4, SPD0280-1/SPD0280-2 and SPD0280-1/SPD0280-2 were used to generate PCR fragments of the left and right flanking regions of *hk08* and *celR*, respectively. PCR products of the erythromycin- and spectinomycin-resistance markers were generated with primers Ery-F/Ery-R and Spec-F/Spec-R-New from plasmid pORI28 and pORI38, respectively (140). Then, by means of overlap extension PCR, the left and right flanking regions of *hk08* and *celR* were fused to these resistance genes. The resulting PCR products were transformed to *S. pneumoniae* D39 wild-type and selection of the mutant strains was done on the appropriate antibiotic. Erythromycin- and spectinomycin-resistant clones were further examined for the presence of the Δ*hk08* and Δ*celR* deletion by PCR.

Table 1: List of strains and plasmids used in this study.

Strain/plasmid	Description[a]	Source
S. pneumoniae		
D39	Serotype 2 strain, *cps 2*	(136)
Δ*celR*	D39 Δ*celR*::Spec^R	This study
Δ*ccpA*	D39 Δ*ccpA*::Spec^R	(43)
Δ*hk08*	D39 Δ*hk08*::Ery^R	This study
SS401	D39 Δ*bgaA*::P*celA-lacZ*; Tet^R	This study
SS402	D39 Δ*bgaA*::P*spd0278-lacZ*; Tet^R	This study
SS403	D39 Δ*bgaA*::P*celB-lacZ*; Tet^R	This study
SS404	D39 Δ*bgaA*::P*celR-lacZ*; Tet^R	This study
SS405	D39 Δ*bgaA*::P*celD-lacZ*; Tet^R	This study
SS406	Δ*celR* Δ*bgaA*::P*celA-lacZ*; Tet^R	This study
SS407	Δ*celR* Δ*bgaA*::P*celB-lacZ*; Tet^R	This study
SS408	Δ*ccpA* Δ*bgaA*::P*celA-lacZ*; Tet^R	This study
SS409	Δ*ccpA* Δ*bgaA*::P*celB-lacZ*; Tet^R	This study
SS410	Δ*hk08* Δ*bgaA*::P*celA-lacZ*; Tet^R	This study
SS411	Δ*hk08* Δ*bgaA*::P*celB-lacZ*; Tet^R	This study
SS412	D39 Δ*bgaA*::P*celA F-lacZ*; Tet^R	This study
SS413	D39 Δ*bgaA*::P*celA H-lacZ*; Tet^R	This study
SS414	D39 Δ*bgaA*::P*celA N-lacZ*; Tet^R	This study
SS415	D39 Δ*bgaA*::P*celB F-lacZ*; Tet^R	This study
SS416	D39 Δ*bgaA*::P*celB H-lacZ*; Tet^R	This study
SS417	D39 Δ*bgaA*::P*celB N-lacZ*; Tet^R	This study
E. coli		
EC1000	Km^R; MC1000 derivative carrying a single copy of the pWV1 *repA* gene in *glgB*	Laboratory collection
Plasmids		
pPP2	Amp^R Tet^R; promoter-less *lacZ*. For replacement of *bgaA* with promoter *lacZ*-fusion. Derivative of pTP1	(88)
pSS401	pPP2 P*celA*	This study
pSS402	pPP2 P*SPD0278*	This study
pSS403	pPP2 P*celB*	This study
pSS404	pPP2 P*celR*	This study
pSS405	pPP2 P*celD*	This study
pSS406	pPP2 P*celA F*	This study
pSS407	pPP2 P*celA H*	This study
pSS408	pPP2 P*celB N*	This study
pSS409	pPP2 P*celB F*	This study
pSS410	pPP2 P*celB H*	This study
pSS411	pPP2 P*celB N*	This study

Construction of *lacZ*-fusions and β-galactosidase assays

Chromosomal transcriptional *lacZ*-fusions to the *celA, celB, celR, celD* and *SPD0278* promoters were constructed in the integration plasmid pPP2 (88) via double crossover in the *bgaA* gene with primer pairs mentioned in Table 2 leading to plasmids pSS401-05. These *lacZ*-fusion constructs were introduced into D39 wild-type resulting in strains SS401-05.

P*celA-lacZ* and P*celB-lacZ* were also transformed to the *celR*, *ccpA* and *hk08* deletion strains resulting in SS406-11. All plasmid constructs were checked by sequencing.

Specific β-galactosidase activity was measured as described before (125) using cells grown in M17 with appropriate sugars (exact concentrations are mentioned in the Result section) which were harvested in the mid-exponential phase of growth.

Table 2: List of primers used in this study.

Name	Nucleotide Sequence (5'→3') [a]	Restriction site
SPD0277-F	CGGAATTCGAATTAGAATTATTGTGAG	EcoRI
SPD0277-R	CGGGATCCGAGCAGCAACAGCACCACC	BamHI
SPD0278-F	CGGAATTCGAAGTTATTTCATCTAACGG	EcoRI
SPD0278-R	CGGGATCCAAAATTGCTGCAATACTTCC	BamHI
SPD0279-F	CGGAATTCGTTTCATTTCTACTCTCAGG	EcoRI
SPD0279-R	CGGGATCCCCTTCAGTAGCACTAATAGC	BamHI
SPD0280-F	CGGAATTCATGGGAATTGAAAAAATGG	EcoRI
SPD0280-R	CGGGATCCAGGATATCTTTCTCTTTAATG	BamHI
SPD0283-F	CGGAATTCCTCACTTCGTTTATAATCC	EcoRI
SPD0283-R	CGGGATCCTTCAAGGAAACCAAACAACC	BamHI
SPD0277-F	GCATGAATTCGAATGCGTTTTTATCTTTTGG	EcoRI
SPD0277-H	GCATGAATTCAATAAGGAAAAAGAATGCG	EcoRI
SPD0277-N	GCATGAATTCTACTTCTTTCCTTAACAATAAG	EcoRI
SPD0279-F	GCATGAATTCTAATTTTTTTCCATATCATTTAGG	EcoRI
SPD0279-H	GCATGAATTCCATTTAGGAAAATGAGGG	EcoRI
SPD0279-N	GCATGAATTCGGGAATTGTCATTGGAAACG	EcoRI
SPD0280-1	TAGTGTTTCATTTCTACTC	-
SPD0280-2	TCCTCCTCACTATTTTGATTAGCCTCCTAAAACGTATGTTTTC	-
SPD0280-3	CCCTTGCATGCATAAACTGCGAATGATGAATTGCAAAATCAGC	-
SPD0280-4	AAGAGTGCTATTGAATAACG	-
Spec-RP-New	GCAGTTTATGCATGCAAGGG	-
Spec-F	CTAATCAAAATAGTGAGGAGG	-
HK08-1	AGTGAGGTTCAATACTTATC	-
HK08-2	GAGATCTAATCGATGCATGCTTGAAATAATATATCCAACC	-
HK08-3	AGTTATCGGCATAATCGTTAACCCTCGTTCTCAACCTCTC	-
HK08-4	CGTTATCCAAACGACGTTCC	-
Ery-F	AACGATTATGCCGATAACT	-
Ery-R	CATGCATCGATTAGATCT	-

[a]Restriction enzyme sites are underlined.

Construction of *celA* and *celB* promoter subclones in pPP2

The following promoter subclones of the *celA* and *celB* promoters were made in pPP2 (88) (primers pairs are mentioned in Table 2): P*celA-F* (truncated just a few bases upstream of the CelR regulatory site), P*celA-H* (with the deletion of half of the CelR regulatory site), P*celA-N* (full CelR regulatory site deleted), P*celB-F* (truncated just a few bases upstream of CelR regulatory site), P*celB-H* (with the deletion of half of the CelR regulatory site) and P*celB-N* (full CelR regulatory site deleted) resulting in plasmid pSS406-11. These *lacZ*-fusion constructs were introduced into D39 wild-type resulting in strains SS412-17.

Microarray analysis

For transcriptome analysis the *S. pneumoniae* D39 wild-type strain and its isogenic *celR* mutant were grown in 3 biological replicates in CM17 (0.5% Cellobiose + M17) medium and harvested at the mid-exponential phase of growth. All other procedures regarding the DNA microarray experiment were performed as described before (126, 225).

Microarray Data Analysis

DNA microarray data were analyzed as describe before (126, 224). In short, DNA Microarray slides were scanned with a Genepix 4200 laser scanner at 10 μm resolution. Array Pro 4.5 (Media Cybernetics Inc., Silver Spring, MD) was used to analyze the slides. The *MicroPrep* software package (253) was used to obtain and analyze the microarray data. The expression ratio of the *celR* mutant over the D39 strain was calculated from the measurements of at least 7 spots by Cyber-T (150). For identification of differentially expressed genes, only genes with a Bayesian p-value <0.001 and fold change cut-off of 3 was applied. Microarray data have been submitted to GEO under accession number GSE30891.

Results

Organization of the cellobiose operon in *S. pneumoniae* D39

The *cel* locus of *S. pneumoniae* D39 wild-type consists of seven genes (*SPD0277-83*), which are organized in two putative transcriptional units starting from P*celA* and P*celB* (Fig. 1). The first transcriptional unit comprises two genes, encoding 6-phospho-β-glucosidase (CelA) and a hypothetical protein (SPD0278), respectively. The second unit encodes the predicted cellobiose-specific PTS components IIBAC (CelB, CelC, and CelD), a DNA binding transcriptional regulator (CelR) and a hypothetical transmembrane protein (SPD0282).

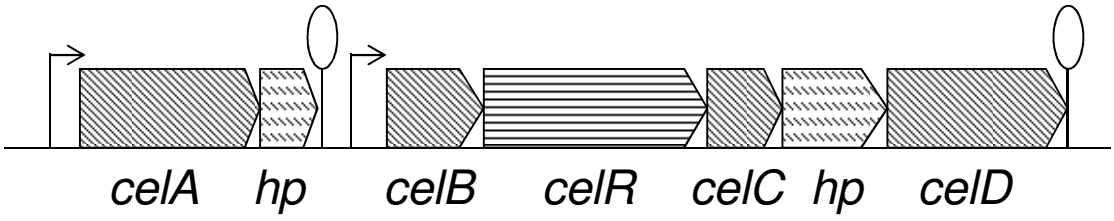


Figure 1: Organization of the *cel* locus in *S. pneumoniae* strain D39. Lollipops and flags indicate putative terminators and putative promoters, respectively.

To prove that the *cel* locus is transcribed into two transcriptional units, transcriptional *lacZ*-fusions to all intergenic regions in this locus were constructed in plasmid pPP2 (88) and

introduced into D39 wild-type. β-galactosidase assays were performed with the strains containing these *lacZ*-fusions grown in CM17 (0.5% Cellobiose + M17). P*celA* and P*celB* were highly active in the presence of cellobiose, while no expression from the other intergenic regions was observed (Table 3). These data demonstrate that the *cel* locus is organized in two transcriptional units originating from the promoters of *celA* and *celB*.

Table 3: Specific β-galactosidase activity (miller units) of D39 wild-type containing the P*celA-lacZ*, P*celB-lacZ*, P*celD-lacZ*, P*celR-lacZ*, and P*SPD0278-lacZ* transcriptional fusions grown in CM17 (0.5% Cellobiose + M17) medium. Standard deviation of 3 independent experiments is given in parentheses.

Specific β-galactosidase Activity (Miller Units)	
Promoters	**CM17**
celA	805 (17)
celB	143 (4.0)
celD	8 (1.1)
celR	5 (1.0)
SPD0278	6 (0.3)

Cellobiose induces and glucose nullifies the expression of the *cel* locus

Table 4: Specific β-galactosidase activity (miller units) of D39 wild-type containing the P*celA-lacZ* or P*celB-lacZ* transcriptional fusions grown in M17 medium supplemented with different concentrations (% w/v) of cellobiose (C) and/or glucose (G). Standard deviation of three independent measurements is given in parentheses.

Specific β-galactosidase Activity (Miller Units)		
Medium	**D39**	
	P*celA*	**P*celB***
M17	80 (3.0)	56 (5.0)
M17+0.2% G	8 (1.0)	6 (0.5)
M17+0.2% C	452 (21)	98(4.0)
M17+0.5% G	7 (1.0)	6 (0.6)
M17+0.5% C	770 (54)	140 (4.0)
M17+0.1% G + 0.1% C	9 (0.9)	7 (1)
M17+0.1% G + 0.2% C	11 (1.2)	9 (0.8)
M17+0.1% G + 0.5% C	10 (1.1)	8 (1)

To investigate the transcriptional response of the *cel* locus of *S. pneumoniae* to different concentrations of cellobiose and glucose, D39 wild-type containing P*celA-lacZ* or P*celB-lacZ* was grown in M17 with different concentrations of these sugars and β-

galactosidase assays were performed with these strains. The data revealed that increasing concentrations of cellobiose lead to increasing expression of both promoters. Increasing the concentration of glucose led to decreasing expression of both promoters (Table 4). To study in more detail the role of cellobiose and glucose in the regulation of the *cel* locus, β-galactosidase activity was measured in D39 containing P*celA-lacZ* and P*celB-lacZ* at a constant concentration of glucose, and with increasing concentrations of cellobiose. Even with a 5 times higher concentration of cellobiose than glucose, the latter prevents transcription from P*celA* and P*celB*. These results show that cellobiose activates the expression of the *cel* locus, while glucose decreases its expression (Table 4).

Role of HK08 and CcpA in the expression of the *cel* operon

In a previous study (168) it was observed that TCS08 is involved in the transcriptional regulation of the *cel* locus in *S. pneumoniae* R6. Therefore, to find out a putative role of HK08 in the sugar-dependent transcription of the *cel* locus in D39, expression of the P*celA-lacZ* and P*celB-lacZ* transcriptional fusions was measured in a deletion strain of *hk08* that was created by allelic replacement with an erythromycin-resistance marker. Interestingly, despite deletion of *hk08*, P*celA* and P*celB* were, in the presence of cellobiose, still expressed to a similar extent as in D39 wild-type, and hardly expressed in the presence of glucose (Table 5). This shows that the observed carbon source-dependent regulation does not go via TCS08, but that TCS08 probably regulates the *cel* locus via a yet to be identified signal.

Table 5: Specific β-galactosidase activity (miller units) of D39 wild-type and the Δ*celR,* Δ*hk08* and Δ*ccpA* mutants all containing the P*celA-lacZ* or P*celB-lacZ* transcriptional fusions grown in M17 medium supplemented with 0.5% (w/v) concentrations of cellobiose (C) or glucose (G). Standard deviation of three independent measurements is given in parentheses.

Specific β-galactosidase Activity (Miller Units)				
	wt	**Δ*hk08***	**Δ*celR***	**Δ*ccpA***
P*celA* (GM17)	7.5 (0.5)	6 (0.1)	6 (0.2)	5 (0.4)
P*celA* (CM17)	791 (27)	801 (33)	5 (0.1)	779 (35)
P*celB* (GM17)	8 (0.2)	7 (0.4)	4 (0.7)	5 (0.2)
P*celB* (CM17)	133 (4.0)	128 (7.0)	5 (0.5)	121 (2.0)

CcpA is responsible for the main part of carbon catabolite regulation in Gram-positive bacteria including *S. pneumoniae*. CcpA mediates its repression by binding to catabolite-responsive element (*cre*) sites present in its target promoters (154, 281). To find out the

putative role of CcpA in the regulation of the *cel* locus, P*celA* and P*celB* were analyzed for the presence of a putative *cre* site. Interestingly, P*celA* contains a putative *cre* box (5'-AAGAATGCGTTTTCAT-3'). However, in the presence of glucose or cellobiose we could not demonstrate an effect of CcpA on the expression of P*celA* and P*celB* (Table 5). Therefore, it is likely that CcpA has, like in *S. mutans* (275), no role in the regulation of the *cel* locus in *S. pneumoniae* and that the *cre* box present in P*celA* is not functional, which is in agreement with a recent study (43).

CelR acts as a transcriptional activator of the *cel* locus in the presence of cellobiose in *S. pneumoniae*

As the *ccpA* and *hk08* mutations have no effect on the regulation of the *cel* locus by cellobiose and glucose, we investigated the role of *celR*, the second gene of transcriptional unit 2 of the *cel* locus, which encodes a putative DNA binding transcriptional regulator. To find the effect of its mutation on the expression of the *cel* locus, P*celA-lacZ* and P*celB-lacZ* transcriptional fusions were introduced in a Δ*celR* strain. Both promoters completely lost their activity even in the presence of cellobiose (Table 5). These data suggest that *celR* acts as a transcriptional activator of the *cel* locus in the presence of cellobiose.

DNA microarray analysis of Δ*celR*

To investigate the effect of the *celR* deletion on the transcriptome of *S. pneumoniae*, D39 wild-type was compared to its isogenic *celR* deletion mutant grown in CM17 (0.5% Cellobiose + M17) medium. CM17 was used for growing the cultures, because as the *celR* mutant grows slower than the wild-type in CM17 medium, cells were harvested at the mid-exponential growth phase according to their growth rate. Table 6 summarizes the results of the transcriptome changes induced in *S. pneumoniae* by the deletion of *celR*. There was no broad effect of the *celR* deletion on the transcriptome of *S. pneumoniae* D39, since after applying the criteria of $\geq$ 3.0-fold difference as the threshold change and a *P* value < 0.001, the *cel* locus was the only cluster of genes that was significantly downregulated in the *celR* deletion strain and no other big responses were observed in the transcriptome. This suggests that the *cel* locus is the only target of CelR that is activated in the presence of cellobiose.

Prediction of a *celR* regulatory site in P*celA* and P*celB*

By applying a MEME motif sampler search (17) a 22-bp palindromic sequence (5'-YTTTCCWTAWCAWTWAGGAAAA-3') located just upstream of *celA* and *celB* was found

in the promoters of *celA* and *celB* of *S. pneumoniae* D39 wild-type (Fig. 2A). This stretch of DNA might serve as a CelR regulatory site in *S. pneumoniae*. The site is conserved in other sequenced strains of *S. pneumoniae* that possess the *cel* locus. *S. mutans, S. gordonii*, *S. dysgalactiae*, *S. gallolyticus*, *S. pyogenes* and *S. suis* contain a *cel* locus with a similar composition as in *S. pneumoniae* as well.

Table 6: Summary of transcriptome comparison of *S. pneumoniae* strain D39 Δ*celR* and D39 wild-type grown in CM17 (0.5% Cellobiose + M17). [a]Ratios >3.0 or <-3.0 (Signal intensity of D39 Δ*celR* divided by that of D39 wild-type). P values all < 0.001.

D39 locus tag	Function	[a]Ratio
SPD0277	6-phospho-β-glucosidase, CelA	-81.2
SPD0278	Hypothetical protein	-2.4
SPD0279	Cellobiose phosphotransferase system IIB component, CelB	-22.5
SPD0280	DNA binding transcriptional regulator, CelR	-13.0
SPD0281	Cellobiose phosphotransferase system IIA component, CelC	-22.8
SPD0282	Hypothetical protein	-14.1
SPD0283	Cellobiose phosphotransferase system IIC component, CelD	-18.8

To find out whether the CelR regulatory site is also conserved in these streptococci the *celA* and *celB* promoters of these species were analyzed for the presence of the putative CelR regulatory sequence. This showed that the CelR regulatory sequence is highly conserved in these streptococci (Fig. 2B). However, a difference in the spacing of a single base pair in the middle of the regulatory site was present in *S. mutans* and *S. dysgalactiae.* The predicted CelR regulatory sequences in the *celA* and *celB* promoters of *S. mutans*, *S. gordonii*, *S. dysgalactia*, *S. gallolyticus*, *S. pyogenes* and *S. suis* were aligned with that of *S. pneumoniae*. In this way, a weight matrix of a 23-bp wide putative CelR regulatory site (5'-WTTTCCKTWKNCRATAMGGAAAA-3') was created (Fig. 2B).

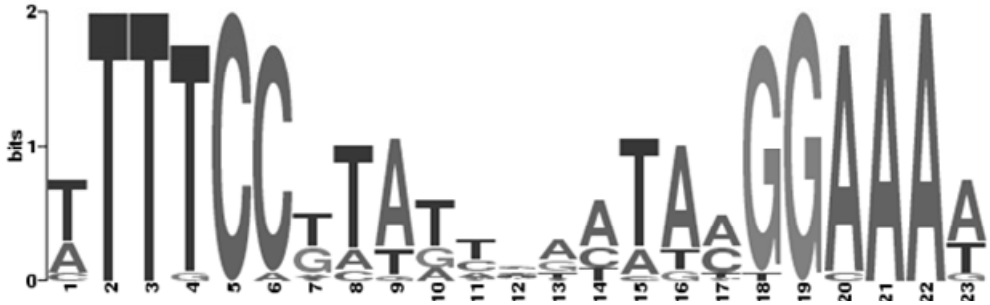
Verification of the CelR regulatory site in P*celA* and P*celB*

To investigate if there are more putative CelR targets in the D39 genome, a genome-wide search with the pneumococcal CelR regulatory site was performed. No other stretch of DNA resembling the CelR regulatory site was found, supporting the observation in the microarray analysis with the *celR* mutant that the *cel* locus is the only target of CelR.

To prove the functionality of the identified CelR regulatory sequence in the regulation of the *cel* locus by CelR, transcriptional *lacZ*-fusions to 5' truncations of P*celA* and P*celB* were constructed (Fig. 3). No expression of both promoters (P*celA* and P*celB*) was observed when the half or entire predicted CelR regulatory site was removed, both in CM17 (0.5% Cellobiose + M17) and GM17 (0.5% Glucose + M17) medium. However, when the *celA* and

celB promoters were truncated only a few base pairs upstream of the predicted regulatory site, expression was similar to that of the full-length promoters (Fig. 3). Therefore, these data suggest that the putative predicted regulatory site is functional and acts as a CelR regulatory site. Moreover, since this predicted site is found highly conserved in other streptococci, it is most likely also functional as a CelR regulatory site in these species.

A



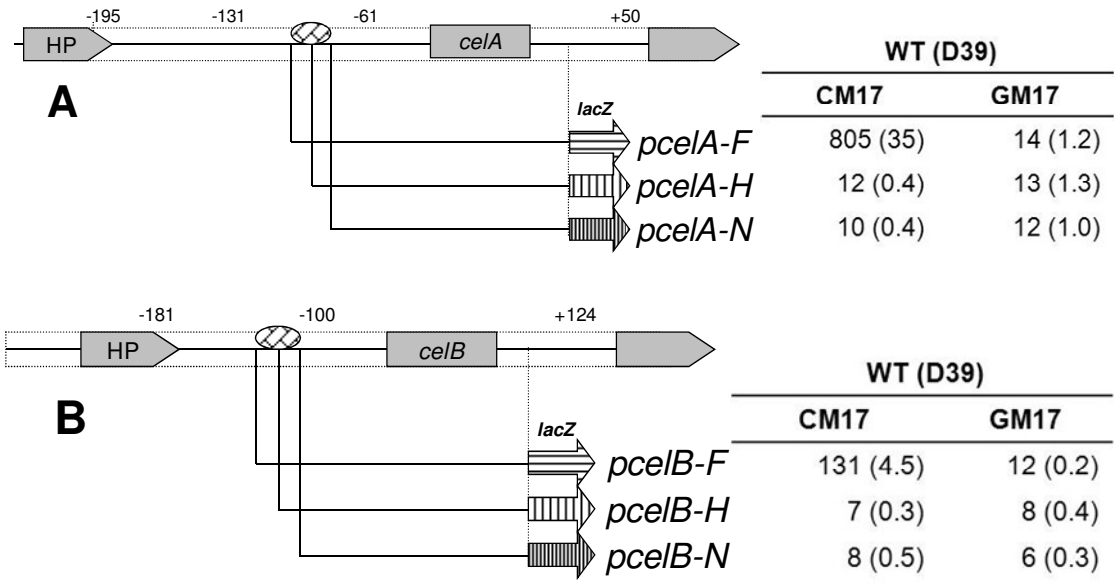
B

1-- TTTCCATTAAAAATTACTT**CTTTCCTTAA CAATAAGGAAAA**AGAATGCGTTTTTATCTTTTGGTTTAA
2-- ATTGAAGTATAAATAATTT**TTTTCCATAT CATTTAGGAAAA**TGAGGGAATTGTCATTGGAAACGCTTT
3-- TAGAATAAAATAAAAGTCT**TTTTCCGTTT CGATATGGAAAA**TCTTGCTATTCAAAACAGCGCTTTCTT
4-- AAGTAGCTTCTTAGAAACC**ATTTCCGTTT CGATACGGAAAA**AGCTCAAATGGAGATTGATGGCGTTTT
5-- AGGCTGCCCTAAGATGAGA**ATTTCCTTAGTCTTTAAGGAAAT**TTTTTCTATTTTAATTTTGAAATTGTT
6-- TAAAGAGGCTGTGTAAGGA**CTTTCCGTTTTCAATACGGAAAA**ACTGTGCGTTTACATTTTGTAAGCGCT
7-- AGAAATAAGAAATCGAGAA**ATTTCCTTAGTCACTAAGGAAAT**TTTTTATATGGTCTATTGTGAAATCGC
8-- TAAAAATTAAATAGATAAT**TTTGCCGTAG TGCTAAGGAAAA**AGTCTCTTTTGAAGTTGTAAAGGTTTT
9-- ATAAATAAAGATGTCAAAA**ATTTCCTAAT AACTTAGGAAAT**TTTTTCTATTCTCGTTTTTGTAACCGC
10-- TTCCGCACTTGGTGAAGTA**TTTTCCTTAT CAAAGAGGAAAA**AGAGTACATTTACTTTATGTAAGCGTT
11-- TCAAGCTTTAAATCACTTA**TTTTCCGTGA TGATAAGGAAAA**AACAGATTTTTTAAAGCGTTTACATCA
12-- ATTGCCCCATCATAGATTC**TTTTCACAAT CAATGCGGAAAA**TGATGGTCACCTTTTTCAGGAAATATT
13-- GAATGATCATGTTGTCTTA**ATTTCCGTTG TGATACGGAAAA**AGAGGGTCTGTGTATCCGTTTACAATA

WTTTCCKTWTTCRATAAGGAAAA

1-- GGTTTATAATAAGAGTATCAAAG-N08-AAAAAAG***ATG*** --- ***celA*** **PN**
2-- ATTTTATAATAAAAATAGATAAC-N47-AAATATT***ATG*** --- ***celB*** **PN**
3-- TTCCTATAATGGAAGTAGTAAAA-N33-ATACAGC***ATG*** --- ***celA*** **SG**
4-- CATATATAATTAAGTTATCAAAT-N18-AATGATC***ATG*** --- ***celB*** **SG**
5-- TGCTAAAATGAAAATAAGAAAAAGGAGGTAGAAGT***ATG*** --- ***celA*** **SM**
6-- TTTTTATAATGGGTATATAAGAA-N14-GAATATT***ATG*** --- ***celB*** **SM**
7-- TTGCTAAAATAAAGTTATAATGA-N10-GGAATTT***ATG*** --- ***celA*** **SD**
8-- CTATTAGAATAAAACCATAAAGAAAAAAGAAATAT***TTG*** --- ***celB*** **SD**
9-- TTGCTATAATATTCTTGTATTAA-N09-TTATGTT***ATG*** --- ***celA*** **Sg**
10-- TTTATAAAATGAACGTGTAAATA-N15-TCACATT***ATG*** --- ***celB*** **Sg**
11-- CCTGTTTTTCTTCTTATAATAAA-N24-AAAGGAT***ATG*** --- ***celA*** **SP**
12-- TGGTACACTAAAGGAGGGTAGAA-N22-AGGACTA***ATG*** --- ***celB*** **SP**
13-- GATGTTCACTGAAATATAATAAG-N23-GAAAGCG***TGT*** --- ***celA*** **SS**

Figure 2: Identification of a CelR regulatory site in different streptococci. **(A)** Weight matrix of the identified CelR regulatory site in P*celA* and P*celB* from different streptococci. **(B)** Position of a CelR regulatory site in P*celA* and P*celB* of different streptococci. PN= *S. pneumoniae*, SM= *S. mutans*, SG= *S. gordonii*, SD= *S. dysgalactiae* Sg= *S. gallolyticus*, SP= *S. pyogenes* and SS= *S. Suis*. Core promoter sequences are underlined, translational start sites are bold-italic and putative CelR regulatory sites are in bold.



	WT (D39)	
	CM17	GM17
pcelA-F	805 (35)	14 (1.2)
pcelA-H	12 (0.4)	13 (1.3)
pcelA-N	10 (0.4)	12 (1.0)

	WT (D39)	
	CM17	GM17
pcelB-F	131 (4.5)	12 (0.2)
pcelB-H	7 (0.3)	8 (0.4)
pcelB-N	8 (0.5)	6 (0.3)

Figure 3: Deletion analysis of P*celA* and P*celB*. A schematic drawing of the P*celA* **(A)** and P*celB* **(B)** truncations is shown. Oval indicates the position of the CelR regulatory site. The tables on the right side give the specific β-galactosidase activity (miller units) of the promoter truncations in D39 wild-type and the *celR* mutant in GM17 (0.5% Glucose + M17) and CM17 (0.5% Cellobiose + M17) medium. Standard deviation of three independent measurements is given in parentheses.

Discussion

In this study, the expression of the *cel* locus of *S. pneumoniae*, which is important for growth on cellobiose, was shown to be activated by cellobiose and repressed by glucose via the transcriptional activator CelR. This regulatory effect depends on a 22-bp palindromic sequence present in the two promoters, P*celA* and P*celB* that drive expression of the *cel* locus.

Expression of the appropriate genes necessary to make efficient use of the available energy sources is of utmost importance for the fitness of pathogenic bacteria in a fluctuating environment. In low GC Gram-positive bacteria, including *S. pneumoniae*, Carbon Catabolite Protein A (CcpA) is mainly responsible for the regulation of sugar metabolic genes in response to the available carbon source(s), which ensures their optimal use (43, 154, 281). In this study, we observed that CcpA plays no role in the regulation of the *cel* locus in *S. pneumoniae*, despite the presence of a putative *cre* site in the promoter of P*celA*. Also, there was no effect on the expression of the *cel* locus in a recent transcriptome-wide analysis of a *ccpA* deletion mutant in glucose and galactose (43). Similarly, deletion of *ccpA* has no impact on the regulation of the *cel* locus in *S. mutans* (275). This means that most probably the *cre* box present in P*celA* is not functional.

A previous study showed that TCS08 activates the *cel* locus in strain R6 (168). We did not find an effect of a deletion of *hk08* on the expression of P*celA* and P*celB* in strain D39. This means that TCS08 probably regulates the *cel* locus in response to a different stimulus than the sugar source in strain D39. Thus, it is likely that CelR is the only regulator involved in the regulation of the *cel* locus in response to the sugar source in strain D39. In addition, our bioinformatics analysis showed that 50% of the strains available at the KEGG website do not contain genes homologous to the *cel* locus, implying that the role of TCS08 lies outside regulation of the *cel* locus in these strains. The absence of the *cel* locus in these strains also suggests the presence of another alternative cellobiose utilization system, which, as determined by BLAST analyses, could be *SPD1830-33*, encoding cellobiose-specific PTS. This second cellobiose system is highly conserved in all the strains available on the KEGG website. Therefore, it will be interesting to determine the role of this system in cellobiose metabolism in future studies.

Bioinformatics analysis revealed that CelR has two PRDs (PTS regulation domain), a HTH (helix-turn-helix), and EIIA and EIIB domains. Usually, transcriptional regulators with PRD domains require phosphorylation on conserved histidine residues by one of the PTS components involved in the metabolism of a certain carbohydrate (256). Recently, the role of CelR (CelR^{SM}) was established in the regulation of the *cel* locus of *S. mutans* (275). CelR^{SM}-dependent regulation of the *cel* locus in the presence of cellobiose required the phosphorylation of CelR^{SM} by CelD. Deletion of the second or both PRD domains present in the CelR^{SM} protein led to the total inactivation of expression of the *cel* locus (275). It was further shown that phosphorylation at H226, H332 and H576 is required for the activation of CelR^{SM} in the presence of cellobiose while phosphorylation of H284 and H391 leads to inhibition of the CelR^{SM}-dependent activation in *S. mutans* in the presence of glucose. Interestingly, alignment of CelR^{SM} with CelR showed that all the above-mentioned histidine residues are conserved in the CelR protein (data not shown). The presence of these conserved histidines suggests that the activity of CelR is determined in the same manner as CelR^{SM}. This is supported by a previous study, showing that *S. pneumoniae* R6 lacking *celD* is unable to grow on cellobiose (168).

Previous STM screens highlight the importance of CelR for the virulence of *S. pneumoniae* (93, 137). So far, however, a possible role in virulence of the other genes in the *cel* locus has never been investigated. Since CelR is dedicated exclusively to the regulation of the *cel* locus, it is likely that the attenuated virulence of a *celR* mutant is caused by de-activation of the expression of the genes present in the *cel* locus. However, since several

pneumococcal strains, like AP200, ATCC700669-23F, G54-19F, Hungary-19A, JJA, P1031 and TCH8431/19A lack the *cel* locus, the contribution of the *cel* locus to virulence seems strain-specific.

As cellobiose is unlikely to be present in the human host, the natural substrate(s) for the PTS encoded by the *cel* locus remain(s) to be determined. It is known that *S. pneumoniae* cleaves host glucosaminoglycans by the activity of many glucosidases, thereby releasing usable carbon sources (121). On the basis of disaccharide units and linkage pattern GAGs are classified into three structural groups (144, 222): (i) the cellobiose type (HA: hyaluronic acid or hyaluronate); (ii) the polylactose type (chondroitin and keratan sulfates); and (iii) the polymaltose type (heparan sulfate) (144). In a previous study (168), it was shown that deletion of *celR* has no effect on the growth of strain R6 in the presence of HA. This suggests that HA is not metabolized via the Cel system and that HA is not involved in the regulation of the *cel* locus. Therefore, it could be that β-linked disaccharides derived from degradation of other host GAGs serve as substrates for the Cel system.

Acknowledgements

Funds for SS were made available under the HEC-NUFFIC overseas scholarship program by HEC (Higher Education commission) Pakistan. TGK is supported by the EU-funded Pneumopath project HEALTH-F3-2009-222983.

Chapter 5

Cellobiose-mediated gene expression in *Streptococcus pneumoniae* and the repressor function of the novel GntR-type regulator CelR-II

Sulman Shafeeq, Oscar P. Kuipers and Tomas G. Kloosterman

(*In preparation*)

Abstract

The human pathogen *Streptococcus pneumoniae* has the ability to use the carbon- and energy source cellobiose due to the presence of a cellobiose-utilizing gene cluster (*cel* locus) in its genome. This system is regulated by the cellobiose-dependent transcriptional activator CelR, which has been previously shown to contribute to pneumococcal virulence. To get a broader understanding of the response of *S. pneumoniae* to cellobiose, we compared the pneumococcal transcriptome during growth on glucose as the main carbon source to that with cellobiose as the main carbon source. The expression of various carbon metabolic genes was altered, including a PTS operon (which we here denote as the *cel-II* operon) that has high homology with the *cel* locus. In contrast to the *cel* locus, the *cel-II* operon is highly conserved in all sequenced strains of *S. pneumoniae*, indicating an important physiological function in the lifestyle of pneumococci. We next characterized the transcriptional regulation of the *cel-II* operon in more detail. Its expression was increased in the presence of cellobiose, while decreased in the presence of glucose. A novel GntR-type transcriptional regulator (which we here denote as CelR-II) was shown to act as a transcriptional repressor of the *cel-II* operon and its repression was relieved in the presence of cellobiose. CelR-II-dependent repression is mediated by a 20-bp DNA operator site (5'-AAAAATGTCTAGACAAATTT-3') present in P*celA-II* that is verified by promoter truncation experiments. In conclusion, we have identified a new cellobiose-responsive PTS operon, together with it transcriptional regulator in *S. pneumoniae*.

Introduction

Streptococcus pneumoniae is a Gram-positive commensal bacterial pathogen in humans that has the ability to colonize the nasopharyngeal cavity of the nose (31). In favorable environmental conditions, it may spread to different parts of the human body to cause serious infections like pneumonia, otitis media, septicemia or meningitis (114, 182) leading to millions of deaths each year, especially in children and the elderly (167). In order to survive in the different niches in the human host, *S. pneumoniae* must have the ability to adapt to fluctuating levels of nutrients (39, 172), such as the available carbohydrate and energy sources.

Previous studies have shown that *S. pneumoniae* is able to use an exceptionally broad spectrum of different carbon sources like cellobiose, raffinose, sucrose, galactose, maltose and others (43, 77, 100, 104, 136, 164, 168, 214, 223). Cellobiose is a β-glucoside carbohydrate

that can also be utilized by *S. pneumoniae* as an energy source alternative to a preferred sugar like glucose (168), although it is not clear whether it encounters cellobiose in its natural environment. However, the extracellular matrix of mammalian tissues is rich in glycosaminoglycans that contain repeating units of β-linked disaccharides (119). The degradation of glycosaminoglycans from the mammalian extracellular tissues may release structural analogues of cellobiose (119, 121). Therefore, a system previously described to function in cellobiose utilization in *S. pneumoniae*, *i.e.* the Cel system (168, 223), may be involved in the acquisition and metabolism of β-glucosides related to cellobiose (29), which are derived by degradation of the extracellular matrix or other biopolymers in the host (119, 121).

The Cel system (encoded by the *cel* locus) (168) has been shown to be required for *S. pneumoniae* strain R6 to grow on cellobiose as the sole energy source. The *cel* loci of strains R6/D39 consist of seven genes and they are transcribed into two transcriptional units (168, 223). The genes in the *cel* locus encode a phospho-β-glucosidase (*celA*), a DNA binding transcriptional activator (*celR*), cellobiose-specific PTS IIBAC components (*celB*, *celC* and *celD*) and two proteins with unknown functions (223). CelR was found to be responsible for activation of the *cel* locus in *Streptococcus mutans* in the presence of cellobiose (275). Recently, CelR of *S. pneumoniae* has been characterized as a transcriptional activator of the *cel* locus as well (223). However, the *cel* locus is not conserved in all pneumococcal strains, as it is absent in 50% of the sequenced *S. pneumoniae* genomes available on the KEGG website, including major multi-drug resistant pneumococcal strains like 19F and 23F (164, 168, 223). Based on this observation, we hypothesized that there might be alternative systems involved in the utilization of β-glucosides as well. To get a broader understanding of the response of *S. pneumoniae* to cellobiose, we looked on a transcriptome-wide level at cellobiose-dependent regulation and found another putative cellobiose/β-glucoside utilization operon, homologous to the *cel* locus, to be highly expressed in the presence of cellobiose. In subsequent experiments this operon was found to be regulated by a GntR-type regulator (CelR-II) encoded by the divergently orientated upstream gene. Expression mediated by CelR-II was increased in the presence of cellobiose and repressed in the presence of glucose. Furthermore, a conserved operator sequence was found that is necessary for the regulation to take place. Since this operon is more conserved than the *cel* locus (see above) among pneumococcal strains, it might be that it has a more important role in utilization of β-linked sugars analogous to cellobiose. However, the exact role of this 'new' operon, which we tentatively name the *cel-II* operon, remains elusive, since in the conditions tested it did not

contribute to growth in medium with cellobiose as the sole carbon source, whereas the *cel* locus did. Therefore, the *cel* locus seems to be the primary transport system for cellobiose.

Material and Methods

DNA manipulation, bacterial strains and growth conditions

Chromosomal DNA of *S. pneumoniae* D39 wild-type (136) was used for PCR amplification. Primers were based on the sequence of the D39 genome (136) and are listed in Table 1.

Table 1: List of primers used in this study.

Name	Nucleotide Sequence (5'→3')[a]	Restriction site
PcelA-II-Fr	CGGGATCCCCGCTAGAAGCTGCTCCCCACC	EcoRI
PcelA-II-Rv	CGGAATTCCTTTTCACGAATCTCATTGT	BamHI
PmalQ-Fr	CGGGAATTCTATGGACGTTTGTGCTTTG	EcoRI
PmalQ-Rv	CGGGATCCGAGATGTGCATCAACACAC	BamHI
PmalP-Fr	CGGGAATTCCCTCTTTAGACAGATTC	EcoRI
PmalP-Rv	CGGGATCCAAGCACCGCAGTGCTC	BamHI
SPD1830-KO-1	GTAAATTCATCACAAGATCC	-
SPD1830-KO-2	TCCTCCTCACTATTTTGATTAGTTTTGTACTCATTTAATCTGG	-
SPD1830-KO-3	CGTTTTAGCGTTTATTTCGTTTAGTCATTACGACATTCCTCCTAGG	-
SPD1830-KO-4	CTGTTTTCATACTCTTTCCC	-
SPD1832-4-KO-1	CTGGATGCCAGACCAATAC	-
SPD1832-4-KO-2	GAGATCTAATCGATGCATGCCAGCAAAGGTGGCAAATTGG	-
SPD1832-4-KO-3	AGTTATCGGCATAATCGTTAGAATTCATCGATCTCTATC	-
SPD1832-4-KO-4	TTCCTGATAGAGTTGTTCAC	-
Spec-R	ACTAAACGAAATAAACGC	-
Spec-F	CTAATCAAAATAGTGAGGAGG	-
Ery-R	TAACGATTATGCCGATAACT	-
Ery-F	GCATGCATCGATTAGATCTC	-
	Forward primers used with PcelA-II-Rv for subclones of P*celA-II*	
PcelA-II-5.3	GCATGAATTCTAGACAAATTTTAAAATTATG	EcoRI
PcelB-II-5.4	GCATGAATTCGTCTAGACAAATTTTAAAATTATGC	EcoRI
PcelB-II-5.5	GCATGAATTCGTATCAACAATTTTTAAAAATG	EcoRI

[a]Restriction enzyme sites are underlined.

Bacterial strains and plasmids used in this study are listed in Table 2. All bacterial strains were stored in 10% (v/v) glycerol at -80 °C. M17 broth (242) supplemented with 0.5% (w/v) glucose was used for growing *S. pneumonia*e D39 wild-type (136) on blood agar plates supplemented with 1% (v/v) defibrinated sheep blood in micro-aerophilic conditions at 37 ^{o}C. For selection, media were supplemented with the following concentrations of antibiotics: erythromycin: 0.25 μg ml^{-1}, spectinomycin: 150 μg ml^{-1}, tetracycline: 2.5 μg ml^{-1} for *S. pneumoniae*; and ampicillin: 100 μg ml^{-1} for *Escherichia coli*.

Table 2: List of strains and plasmids used in this study.

Strain/plasmid	Description	Source
S. pneumoniae		
D39	Serotype 2 strain. *cps 2*	Laboratory of P. Hermans.
ΔccpA	D39 Δ*ccpA*; $Spec^R$	(43)
Δ*celR*	D39 Δ*celR*; $Spec^R$	(223)
ΔcellDBC-II	D39 Δ*celDBC-II*; Ery^R	This study
SS300	D39 Δ*celR-II*; $Spec^R$	This study
SS301	D39 Δ*bgaA*::P*celA-II-lacZ*; Tet^R	This study
SS302	SS300 Δ*bgaA*:: P*celA-II-lacZ*; Tet^R	This study
SS303	*ΔccpA* Δ*bgaA*::P*celA-II-lacZ*; Tet^R	This study
SS304	*ΔcellDBC-II* Δ*bgaA*::P*celA-II-lacZ*; Tet^R	This study
SS305	D39 Δ*bgaA*::P*celA-II-5.3-lacZ*; Tet^R	This study
SS306	D39 Δ*bgaA*::P*celA-II-5.4-lacZ*; Tet^R	This study
SS307	D39 Δ*bgaA*::P*celA-II-5.5-lacZ*; Tet^R	This study
E. coli		
EC1000	Km^R; MC1000 derivative carrying a single copy of the pWV1 *repA* gene in *glgB*	Laboratory collection
Plasmids		
pPP2	Amp^R Tet^R; promoter-less *lacZ*. For replacement of *bgaA* with promoter *lacZ*-fusion. Derivative of pTP1	(88)
pSS301	pPP2 P*celA-II-lacZ*	This study
pSS302	pPP2 P*celA-II-5.3-lacZ*	This study
pSS303	pPP2 P*celA-II-5.4-lacZ*	This study
pSS304	pPP2 P*celA-II-5.5-lacZ*	This study

Construction of deletion mutants of *celR-II* and *celDBC-II*

celR-II and *celDBC-II* deletion mutants were made by allelic replacement with a spectinomycin- and erythromycin-resistance marker, respectively, following the procedure as described before (224). Briefly, primer pairs SPD1830-KO-1/SPD1830-KO-2, SPD1830-KO-3/SPD1830-KO-4, 1832-4-KO-1/1832-4-KO-2 and 1832-4-KO-3/1832-4-KO-4 were used to generate PCR fragments of the left- and right flanking regions of *celR-II* and *celDBC-II*, respectively. The spectinomycin-resistance marker was amplified by a PCR on pORI38 (140) with primers Spec-F/Spec-R. The erythromycin-resistance marker was amplified by a PCR on pORI28 (140) with primers Ery-F/Ery-R. Then, the left- and right flanking regions of *celR-II* and *celDBC-II* were fused to the spectinomycin- and erythromycin-resistance markers, respectively, by means of overlap-extension PCR. The resulting PCR products were transformed to *S. pneumoniae* D39 wild-type and selection of the mutant strains was done with the appropriate antibiotic. Spectinomycin- and erythromycin-resistant clones were further examined for the presence of the *celR-II* and *celDBC-II* deletion by PCR.

Construction of *lacZ*-fusions and *celA-II* promoter subclones in pPP2

The pPP2 (88) plasmid was used to construct a vector for ectopic chromosomal transcriptional *lacZ*-fusions to the promoter of *celA-II* in *S. pneumoniae* D39 using the primer pair mentioned in Table 1. The resulting plasmid pSS301 was introduced into the D39 wild-type and *celR-II* mutant strains, integrated in the genome via double crossover in the *bgaA* gene, resulting in strains SS301 and SS302. The P*celA-II-lacZ* fusion was also introduced into the *ΔccpA* strain that was published before (43) and in the *celDBC-II* mutant, resulting in strain SS303 and SS304, respectively.

The following promoter subclones of P*celA-II* were made in pPP2 (primers pairs are mentioned in Table 1; a schematic picture of the P*celA-II* subclones is shown in Fig. 3): P*celA-II-5.3* (truncated 15 bases upstream of the CelR-II operator site), P*celA-II-5.4* (first 6 bases of the CelR-II operator site deleted) and P*celA-II-5.5* (half of the CelR-II operator site deleted, but keeping the -35 site intact). This resulted in plasmids pSS302-04. These constructs were introduced into D39 wild-type as described above, resulting in strains SS305-07. All plasmid constructs were checked by sequencing.

Enzyme assays

Specific β-galactosidase assays were performed as described before (125). Cultures of different strains were grown in M17 in the presence of the appropriate carbon source (exact concentrations are mentioned in the Result section) and harvested in the mid-exponential phase of growth.

Experimental procedure of DNA microarray analysis

For transcriptome analysis of *S. pneumoniae*, the D39 wild-type strain was grown in 3 biological replicates in CM17 (0.5% Cellobiose + M17) and compared in triplicate to the strain grown in GM17 (0.5% Glucose + M17) medium. Cultures were harvested for RNA isolation at two growth points in CM17 (CT-1 and CT2) and GM17 (GT-1 and GT-2) medium (see Fig. 1). To analyze the effect of *celR-II* deletion on the transcriptome, *S. pneumoniae* D39 wild-type and its isogenic *celR-II* mutant (SS300) were grown in 3 biological replicates in GM17 (0.5% Glucose + M17) medium. These cultures were harvested at the mid-exponential phase of growth at an OD600 of 0.3. All other procedures regarding the DNA microarray experiments (cell disruption, RNA isolation, RNA quality testing, cDNA synthesis, labeling with dyes (Cy3 and Cy5), hybridization and scanning) were performed as described before (126, 225).

DNA microarray data analysis

DNA microarray data were analyzed as describe before (126, 225). For identification of differentially expressed genes a Bayesian p-value <0.001 and fold change cut-off of 3 was applied. Microarray data have been submitted to GEO.

Results

Growth and transcriptome analysis of *S. pneumoniae* in medium with cellobiose as the sole carbon source

To study the growth behavior of *S. pneumoniae* D39 in cellobiose, growth experiments were performed in the nutrient rich M17 medium supplemented with 0.5% cellobiose, which was compared to the growth in M17 + 0.5% glucose. As expected, normal exponential growth of D39 was observed in the presence of glucose. However, during growth of *S. pneumoniae* on cellobiose two distinct exponential growth phases were observed. As the first exponential growth phase is similar to the growth in M17 without added carbon source (Fig. 1), it is likely that cellobiose is only metabolized in the second growth phase. Because of these two growth phases on cellobiose, we decided to perform a transcriptome comparison with the growth on glucose in both phases (T1 and T2, Fig. 1) and in this way identify cellobiose-responsive genes. Therefore, D39 wild-type was grown in CM17 (0.5% Cellobiose + M17) and GM17 (0.5% Glucose + M17), and cells were harvested at the time points as indicated in Fig. 1. Table 3 summarizes the number of genes that were affected in this transcriptome profiling experiment, grouped into COG functional categories on the basis of the putative functions of the corresponding proteins. At T1 >3 times more genes were affected than in T2, comprising many genes of COG J (translation, ribosomal structure and biogenesis), probably reflecting the fact that in the presence of cellobiose cells have to adapt in the first phase and are therefore not growing optimally, whereas in the second phase the carbon source is likely to be actively metabolized. At both time points, most of the affected genes belong to COG functional category G (carbohydrate transport and metabolism). Almost all the affected genes belonging to category G (49 out of 51 at T1 and 16 out of 20 at T2) were upregulated in the presence of cellobiose. These effects may be either due to release of carbon catabolite repression of these genes as an effect of the absence of glucose, or due to a direct inductive effect of cellobiose. Other COG functional categories with a high number of differentially expressed genes are S (function unknown) and R (general functions prediction only).

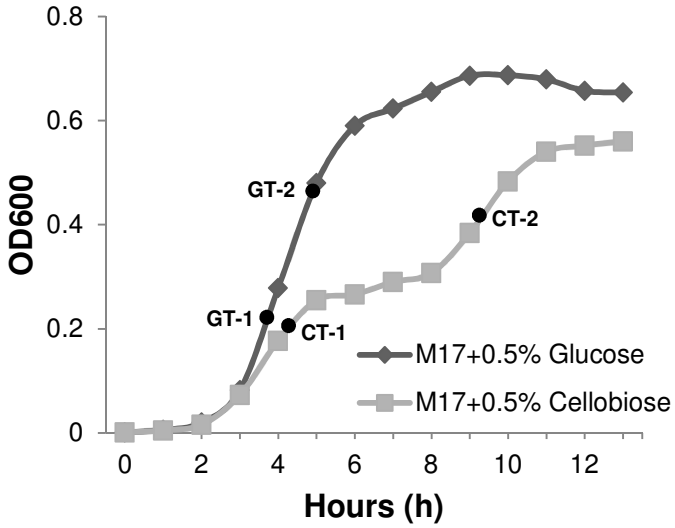


Figure 1: Growth of *S. pneumoniae* D39 in the presence of 0.5% cellobiose (grey line ■) and 0.5% glucose (black line ♦) in M17 medium. Black circles show the time points on which cultures are harvested for transcriptome analysis.

Table 3: Number of genes significantly* affected in the presence of cellobiose at time point T1 and T2.

Functional Categories	**T1**			**T2**		
	Total	**Down**	**Up**	**Total**	**Down**	**Up**
C: Energy production and conversion	7	1	6	7	5	2
D: Cell cycle control, cell division, chromosome partitioning	6	5	1	1	0	1
E: Amino acid transport and metabolism	18	16	2	5	3	2
F: Nucleotide transport and metabolism	14	12	2	4	4	0
G: Carbohydrate transport and metabolism	**51**	**2**	**49**	**20**	**4**	**16**
H: Coenzyme transport and metabolism	9	4	5	2	1	1
I: Lipid transport and metabolism	3	2	1	2	0	2
J: Translation, ribosomal structure and biogenesis	26	25	1	0	0	0
K: Transcription	12	8	4	4	0	4
L: Replication, recombination and repair	14	10	4	2	1	1
M: Cell wall/membrane/envelope biogenesis	14	8	6	4	3	1
O: Posttranslational modification, protein turnover, chaperones	6	5	1	4	0	4
P: Inorganic ion transport and metabolism	8	6	2	2	1	1
Q: Secondary metabolites biosynthesis, transport and catabolism	1	0	1	0	0	0
R: General function prediction only	21	7	14	4	0	4
S: Function unknown	48	23	25	21	4	17
T: Signal transduction mechanisms	2	1	1	0	0	0
U: Intracellular trafficking, secretion, and vesicular transport	3	1	2	0	0	0
V: Defense mechanisms	13	3	10	3	2	1
Total number of genes	**276**	**139**	**137**	**85**	**28**	**57**

* Representing the genes with at least 3 fold increase or 3 fold decrease in expression levels in CM17 compared to GM17, and with a Bayesian p-value below 0.001 (Cyber-T test).

The effects of cellobiose on the transcriptome of *S. pneumoniae* D39 that were observed at both time points are summarized in Table 4. Expression of various genes and operons of diverse functions was altered in the presence of cellobiose. Notably, genes involved in cellobiose and maltose metabolism were highly upregulated. In addition, an ABC

transporter cluster (*msmEFG*) that encodes a putative multiple sugar transport system was highly upregulated as well.

Table 4: List of genes that are differentially expressed in the transcriptome comparison of *S. pneumoniae* D39 strain grown in CM17 and GM17 at both time points. [a]Gene numbers refer to D39 locus tags. [b]D39 annotation/TIGR4 annotation (100, 136, 243), [c]Ratio represents the fold increase in the expression of genes in CM17 as compared to GM17. In some cases neighbouring genes with lower ratios are also indicated.

D39 locus tag[a]	Function (TIGR Annotation)[b]	Ratio[c]	
		T1	T2
SPD0265	Alcohol dehydrogenase, zinc-containing	26.0	2.1
SPD0277	6-Phospho-β-glucosidase, CelA	71.9	88.4
SPD0278	Hypothetical protein	2.1	3.5
SPD0279	Cellobiose phosphotransferase system IIB component, CelB	13.9	23.4
SPD0280	DNA binding transcriptional regulator, CelR	10.3	10.5
SPD0281	Cellobiose phosphotransferase system IIA component, CelC	11.7	6.9
SPD0282	Hypothetical protein	12.8	4.3
SPD0283	Cellobiose phosphotransferase system IIC component, CelD	8.8	3.7
SPD0344	DNA-binding response regulator	3.5	2.8
SPD0466	BlpT protein fusion	4.3	4.0
SPD0473	Immunity protein BlpY	11.8	5.6
SPD0661	PTS system IIABC components	2.0	2.1
SPD0850	Lactoylglutathione lyase, GloA	-20.6	2.5
SPD0851	Dihydroorotate dehydrogenase, PyrK	-150.3	2.8
SPD0852	Dihydroorotate dehydrogenase B, PyrDb	-65.8	2.3
SPD0886	Thioredoxin family protein	12.6	3.6
SPD1495	Sugar ABC transporter, sugar-binding protein	8.5	7.1
SPD1496	PTS system. IIBC components	2.7	2.7
SPD1590	General stress protein 24. putative	5.5	5.1
SPD1675	Sugar ABC transporter, MsmG	8.2	13
SPD1676	Sugar ABC transporter, MsmF	11.9	1.7
SPD1677	Sugar ABC transporter, MsmE	10.7	2.1
SPD1726	Pneumolysin. PlY	-2.0	-2.0
SPD1727	Hypothetical protein	-2.7	-2.5
SPD1728	Hypothetical protein	-3.5	-3.6
SPD1729	Hypothetical protein	-8.64	-3.4
SPD1830	Glycosyl hydrolase family 1, CelA-II	57.1	2.3
SPD1831	PTS system. IIC component, CelD-II	62.9	3.4
SPD1832	PTS system. IIB component, CelB-II	53.4	4.2
SPD1833	PTS system. IIA component, CelC-II	55.9	3.3
SPD1865	Alcohol dehydrogenase. zinc-containing	12.6	2.5
SPD1866	N-Acetylglucosamine-6-phosphate deacetylase, NagA	4.2	2.5
SPD1933	Glycogen phosphorylase family protein	1.3	1.7
SPD1934	4-Alpha-glucanotransferase, MalQ	2.1	3.4
SPD1935	Maltose/maltodextrin ABC transporter, MalX	1.3	2.1
SPD1936	Maltodextrin ABC transporter, MalC	1.6	1.8

Cellobiose metabolism has previously been shown to be carried out by a cellobiose-specific gene cluster (*cel* locus) in streptococci, including *S. pneumoniae* (168, 223, 275). In our previous studies, we showed that CelR is involved in the activation of the *cel* locus, specifically in the presence of cellobiose (223). As expected, the *cel* locus was also among the highly upregulated genes at both time points. This suggests that the conditions applied for the

transcriptome analysis are indeed appropriate to find cellobiose-responsive genes. The expression of an operon (*SPD1830-1833*, which will be denoted here tentatively as the *cel-II* operon) encoding a glycosyl hydrolase protein and the PTS system IICBA components was upregulated at both time points of growth. Interestingly, blast searches showed that this operon has high homology with the *cel* locus that was previously identified to be involved in cellobiose metabolism (168, 223). This finding, together with its upregulation in the presence of cellobiose might indicate a putative role of this *cel-II* operon in metabolism/utilization of cellobiose/similar carbon sources. Therefore, we decided to further investigate the regulation of this operon in the presence of cellobiose.

Organization and conservation of the *cel-II* operon in *S. pneumoniae*

The *cel-II* locus spans the genes *SPD1830-33* (Fig. 2). *SPD1830* (here named *celA-II*) encodes a glycosyl hydrolase belonging to the BglB family which has high homology to *celA* of *S. pneumoniae* and other streptococci. Next, the downstream three genes (*SPD1831-33*, named *celDCB-II*), encode a PTS system IICBA components that show high homology with *celDCB* located in the *cel* locus of various streptococci including *S. pneumoniae*. Upstream of *celA-II*, a gene encoding a GntR family transcriptional factor (named *celR-II*) is located. The presence of this transcription factor in the DNA region upstream of the *cel-II* operon may indicate that it functions as a regulator of the *cel-II* operon (see also below).

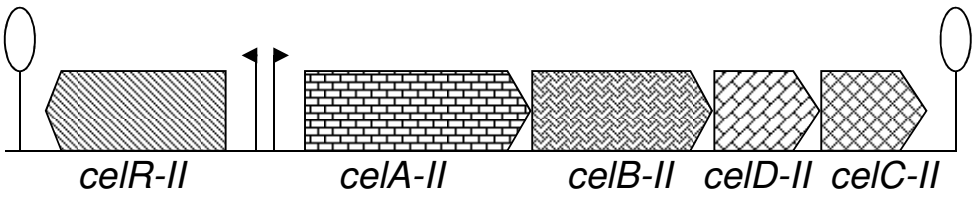


Figure 2: Organization of the *cel-II* operon in *S. pneumoniae* D39. Lollipops indicate the putative transcriptional terminators. Black arrows represent promoter regions. See text for further details.

In a previous study we have studied the role of CelR in the regulation of the *cel* locus (223). However, the *cel* locus was found to be absent from ~50% of the *S. pneumoniae* strains present in the KEGG database. We also analyzed all the strains of *S. pneumoniae* available on the KEGG website for the presence of the *cel-II* operon. Remarkably, this operon is 100% conserved in all the sequenced strains of *S. pneumoniae*. This indicates an important role in carbohydrate metabolism, probably of cellobiose or related/analogous compounds. In subsequent experiments, we focused in more detail on the molecular basis of the transcriptional regulation of the *cel-II* operon.

Cellobiose-dependent expression of the *cel-II* operon

To examine whether the observed altered expression of the *cel-II* operon in the transcriptome analyses is due to a specific effect of cellobiose, we constructed an ectopic transcriptional *lacZ*-fusion to the *celA-II* promoter in the D39 wild-type strain. This strain was grown in M17 medium supplemented with 0.5% of different carbon sources (Cellobiose, Dextrose, Fructose, Galactose, Glucose, Lactose, Mannose, Maltose, NAG, Raffinose, Sucrose and Trehalose). The highest expression of P*celA-II* was observed in the presence of cellobiose (Table 5). The presence of glucose led to the lowest expression of P*celA-II*. These results confirm the microarray data and additionally show that expression of P*celA-II* is controlled specifically by cellobiose.

Table 5: Specific β-galactosidase activity (miller units) of D39 wild-type containing the P*celA-II-lacZ.* grown in M17 medium with different added sugars (0.5% w/v). Standard deviation of 3 independent experiments is given in parentheses.

Specific β-galactosidase Activity (Miller Units) in M17 medium	
Sugars	P*celA-II-lacZ*
No	124 (4)
Cellobiose	515 (6)
Dextrose	108 (5)
Fructose	121 (9)
Galactose	122 (7)
Glucose	105 (6)
Lactose	125 (9)
Mannose	129 (5)
Maltose	131 (6)
NAG	130 (8)
Rafinose	116 (7)
Sucrose	103 (5)
Trehalose	125 (4)

To further investigate the role of glucose in the regulation of P*celA-II*, we grew the cells in the presence of a certain constant concentration of glucose, with an increasing concentration of cellobiose, as mentioned in Table 6. At 0.2% glucose, the lowest expression of P*celA-II* expression was observed. This level of expression could not be increased by adding 0.5% cellobiose to the medium. Thus, cellobiose increases the expression of P*celA-II*, whereas glucose decreases its expression, whereby the last compound appears to overrule the first when present at the same time.

Table 6: Specific β-galactosidase activity (miller units) of D39 wild-type containing the P*celA-II-lacZ.* grown in M17 medium with different combinations of added sugars (% w/v). Standard deviation of 3 independent experiments is given in parentheses. G, Glucose. C, Cellobiose.

Specific β-galactosidase Activity (Miller Units) in M17 medium	
Sugars	P*celA-II-lacZ*
No	115 (2)
0.1% C	205 (3)
0.2% C	340 (5)
0.3% C	380 (5)
0.5% C	490 (17)
1% C	650 (13)
0.1% G	109 (2)
0.2% G	107 (3)
0.3% G	103 (7)
0.5% G	105 (5)
1% G	101 (8)
0.1% G + 0.1% C	112 (6)
0.1% G + 0.2% C	122 (4)
0.1% G + 0.5% C	140 (8)
0.2% G + 0.1% C	109 (1)
0.2% G + 0.2% C	108 (3)
0.2% G + 0.5% C	112 (5)

SPD1829 (*celR-II*) is a repressor of the *cel-II* operon

Adjacent to the *cel-II* operon, a putative GntR-family regulator is located (*SPD1829*), which we named CelR-II. We hypothesized that this regulator could be involved in the observed regulation of the *cel-II* operon. To investigate the role of CelR-II, the *celR-II* gene was replaced by a spectinomycin marker by means of allelic replacement. To examine the effect of the *celR-II* deletion on the transcriptome of strain D39, the transcriptome of the *celR-II* mutant strain (SS300) of *S. pneumoniae* was compared to that of the D39 wild-type strain grown in GM17 (0.5% glucose + M17) medium. GM17 medium was used, since low expression of the *cel-II* operon was observed in the presence of glucose, which we hypothesized to represent a condition with maximal repression of the *cel-II* operon. These transcriptome data revealed that the expression of several genes was altered due the deletion of *celR-II* (Table 7). The expression of *celR-II* was five-fold downregulated confirming the inactivation of the *celR-II* gene in the *celR-II* deletion strain (SS300). The most highly upregulated genes were the ones constituting the *cel-II* operon, indicating a role/function of

CelR-II as the repressor of the *cel-II* operon. Another strong effect caused by deletion of *celR-II* was the upregulation of an operon involved in maltose/maltodextrin metabolism (179). These results show that inactivation of *celR-II* brings about only a modest effect on the transcriptome of *S. pneumoniae* D39, and furthermore imply a role of CelR-II in repressing the *cel-II* operon.

Table 7: Summary of transcriptome comparison of *S. pneumoniae* strain D39 Δ*celR-II* and D39 wild-type grown in GM17. [a]Gene numbers refer to D39 locus tags. [b]D39 annotation/TIGR4 annotation (100, 136, 243), [c]Ratio represents the fold increase in the expression of genes in CM17 as compared to GM17.

D39 locus tag[a]	Function[b]	Ratio[c]
SPD0311	Glucan 1,6-alpha-glucosidase, DexB	3.3
SPD0661	PTS system, IIABC components	2.9
SPD0771	Lactose phosphotransferase system repressor, LacR	4.6
SPD0772	1-phosphofructokinase, putative	4.7
SPD0773	PTS system, fructose specific IIABC components	5.2
SPD1830	Glycosyl hydrolase, family 1, CelA-II	96.3
SPD1831	PTS system, IIC component, CelD-II	129.0
SPD1832	PTS system, IIB component, CelB-II	49.2
SPD1833	PTS system, IIA component, CelC-II	26.7
SPD1932	Glycogen phosphorylase family protein	10.0
SPD1933	4-alpha-glucanotransferase, MalQ	14.7
SPD1934	Maltose/maltodextrin ABC transporter, MalX	4.7
SPD1935	Maltodextrin ABC transporter, MalC	2.9
SPD0593	Elongation factor Tu family protein	-14.0
SPD1829	GntR family transcriptional regulator, CelR-II	-5.3

Regulation of P*celA-II* in *ccpA*, *celR-II* and *celDBC-II* mutant strains

To further confirm that CelR-II is responsible for the repression of the *cel-II* operon that was upregulated in the *celR-II* mutant strain (SS300), we introduced the transcriptional *lacZ*-fusion to P*celA-II* into the D39 wild-type and *celR-II* mutant strains. In GM17 (0.5% Glucose + M17) medium specific β-galactosidase activity was highly increased in the *celR-II* mutant as compared to D39 wild-type (Table 8). This not only confirms the data of the *celR-II* mutant transcriptome but also suggests that CelR-II carries out repression via the P*celA-II* promoter.

CcpA is considered a master transcriptional regulator in the control of carbohydrate utilization and metabolism genes in Gram-positive bacteria including *S. pneumoniae* (43, 154, 281). To investigate a possible role of CcpA in the regulation of the *cel-II* operon, we measured specific β-galactosidase activity of P*celA-II-lacZ* in a *ccpA* mutant strain. No difference in expression of P*celA-II* was observed in the *ccpA* mutant as compared to the D39 wild-type strain, when cells were grown in M17 with either cellobiose or glucose (Table 8). In

addition, no CcpA binding site (*cre*) was found in the P*celA-II* promoter. Thus, these data show that regulation of the *cel-II* operon in *S. pneumoniae* is independent of CcpA.

In *S. mutans*, regulation of the *cel* locus by CelR requires the phosphorylation of CelR by one of the PTS components, namely CelD (275). Therefore, to investigate a possible role of the *celDBC-II* mutant in the regulation of the *cel-II* operon via an effect on CelR-II activity, we measured specific β-galactosidase activity of P*celA-II-lacZ* in a *celDBC-II* mutant strain. No difference in expression of P*celA-II* was observed in the *celDBC-II* mutant as compared to wild-type D39, when cells were grown in M17 medium with cellobiose or glucose (Table 8). This suggests that the components of the Cel-II PTS system are not required for activation of CelR-II in the presence of cellobiose. However, we cannot exclude that other PTS systems confer a regulatory effect on CelR-II.

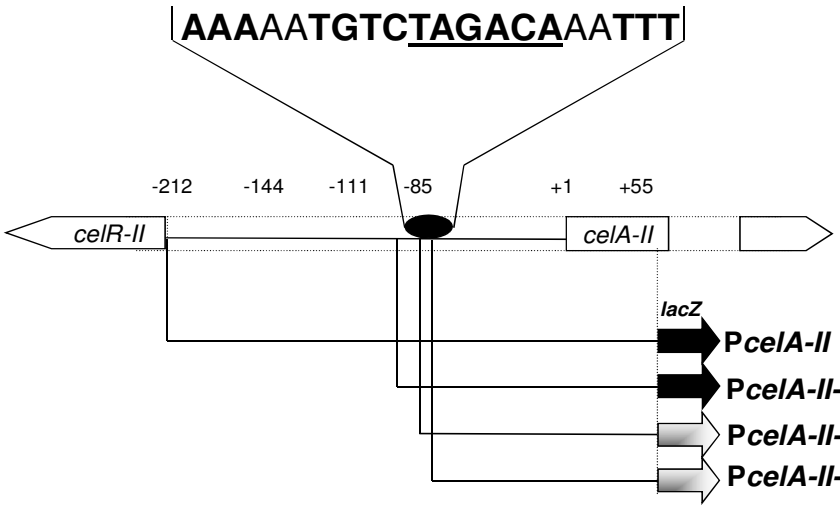
Table 8: Specific β-galactosidase activity (miller units) of D39 wild-type, *ΔcelR-II*, *ΔccpA*, and *ΔcellDBC-II* mutants all containing the P*celA-II-lacZ* transcriptional fusion grown in M17 medium supplemented with added concentrations (0.5% w/v) of cellobiose (C) and glucose (G). Standard deviation of three independent measurements is given in parentheses.

Specific β-galactosidase Activity (Miller Units)				
	WT	***ΔcelR-II***	***ΔccpA***	***ΔcellDBC-II***
P*celA-II* (GM17)	115 (5)	1660 (17)	111 (8)	116 (9)
P*celA-II* (CM17)	610 (10)	1690 (15)	580 (12)	599 (17)

Identification of the CelR-II operator site

The above data strongly suggest a direct effect of CelR-II on P*celA-II*. To identify a possible CelR-II operator sequence, a 5' promoter truncation study was performed with the *celA-II* promoter. A diagram of the *celA-II* promoter truncation is shown in Fig. 3. Truncation of P*celA-II* near to the predicted -35 core promoter sequence (P*celA-II-5.4*) relieved the repressive action of CelR-II on P*celA-II,* suggesting the presence of a putative CelR-II operator in this deleted region in the promoter. Further bioinformatics analysis of this area revealed the presence of a 20-bp palindromic region (5'-AAAAATGTCTAGACAAATTT-3') that is overlapping with the -35 site and that might act as the CelR-II operator site. Deletion of half of this predicted operator site (P*celA-II-5.5*) led to high expression of P*celA*-II in CM17 (0.5% Cellobiose + M17) and GM17 (0.5% Glucose + M17) medium (Fig. 3). However, when P*celA-II* was truncated only a few base pairs upstream of the predicted operator site (P*celA-II-5.3*), expression was similar to that of the full-length promoters (Fig. 3). Therefore,

these data suggest that the putative predicted operator site is functional and acts as the CelR-II operator site.



	GM17	CM17
PcelA-II	116 (3)	615 (13)
PcelA-II-5.3	113 (3)	613(20)
PcelA-II-5.4	1470 (25)	1571 (23)
PcelA-II-5.5	1585 (15)	1580 (17)

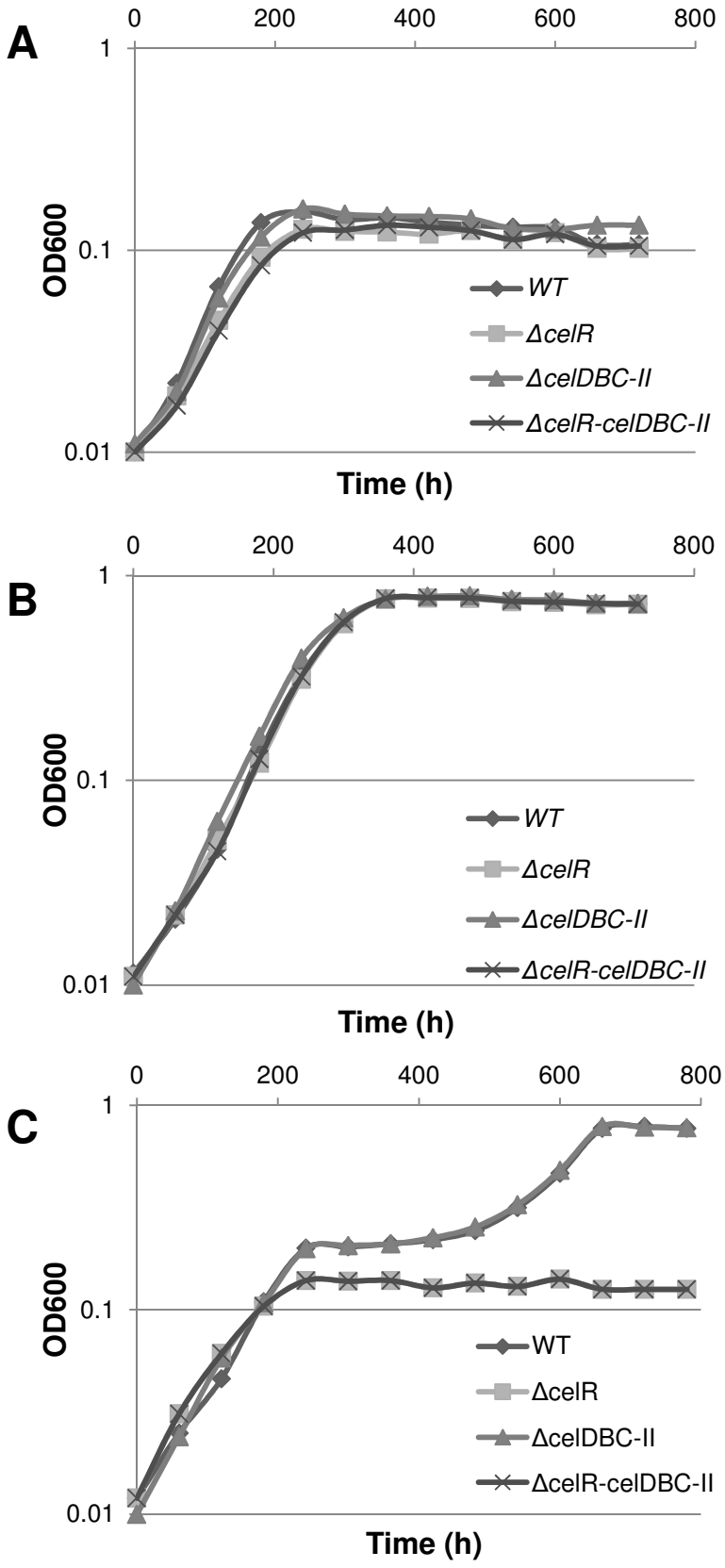
Figure 3: Analysis of truncations of P*celA-II*. A schematic overview of the *celA-II* promoter truncations is shown. The table on the right gives the specific β- galactosidase activity of the truncated promoters in GM17 (0.5% glucose + M17) and CM17 (0.5% cellobiose + M17). The sequence of the predicted CelR-II operator sequence is indicated, with palindromic bases in bold and an underlined -35

The *cel* locus is required for D39 to grow on cellobiose, while the *cel-II* operon is not

As in a previous study (168), it has been shown that deletion of *celR* or *celD* from the *cel* locus leads to growth inhibition in CDM with cellobiose as the carbon source, we decided to mutate *celDBC-II* and compare the growth of *ΔcelDBC-II*, *ΔcelR* (which has no expression of the *cel* locus) (168, 223), and *ΔcelR-celDBC-II* in the presence of 0.5% cellobiose and 0.5% glucose with that of D39 wild-type (Fig. 4). In the presence of glucose, all strains grew similar as the wild-type. However, clear growth differences were observed in the presence of cellobiose. As mentioned before, D39 wild-type grows in two distinct exponential phases in the presence of cellobiose. A similar growth pattern was observed for the *ΔcelDBC-II* mutant strain, but the *ΔcelR* and *ΔcelR-celDBC-II* mutant strains were not able to start the second exponential phase.

These data suggest that the *cel* locus is important for *S. pneumoniae* D39 to grow on cellobiose, while under these conditions the *cel-II* operon, although its expression is responsive to cellobiose, is not.

Figure 4: Growth of *S. pneumoniae* D39 wild-type (♦) and its isogenic mutants *celR* (■), *celDBC-II* (▲) and *celR-celDBC-II* (**X**) in M17 (A), 0.5% glucose + M17 (B) and 0.5% cellobiose + M17 (C).



Discussion

In many bacteria, including *S. pneumoniae*, glucose is considered as one of the primary sources of energy for metabolic processes. However, the existence of numerous other sugar-specific systems in *S. pneumoniae* implies its ability to utilize various other carbon sources in the absence of glucose (29, 100, 136). The role of different systems dedicated to the uptake and metabolism of different sugars including cellobiose, maltose, galactose, sucrose, raffinose and others has been explicitly investigated for *S. pneumoniae* (29, 43, 77, 159, 160, 168, 179, 223, 250). In a previous study, we described the role of the cellobiose-dependent transcriptional activator (CelR) in the regulation of the cellobiose utilization gene cluster (*cel* locus) (223). Intriguingly, the *cel* locus is not conserved in all the sequenced strains of *S. pneumoniae* that are available on the KEGG website, indicating that *S. pneumoniae* harbors other ways of cellobiose utilization. To be able to more thoroughly understand the effect of cellobiose on *S. pneumoniae*, we applied transcriptome profiling. The data show that, besides the *cel* operon, a second operon (*cel-II*) is responsive to cellobiose as well. Interestingly, the fact that the *cel-II* operon is 100% conserved in different pneumococcal strains, suggests that it has an important physiological function in the life style of *S. pneumoniae*. Indeed, like *celR*, *celD-II* was also found in one of the STM studies, where it was implicated in lung infection (93). Furthermore, in this study, we show that the expression of the *cel-II* operon increases in the presence of cellobiose, but decreases in the presence of glucose. Additionally, we identified and characterized the role of a GntR-type transcriptional regulator (CelR-II) in the regulation of the *cel-II* operon.

The phosphoenolpyruvate-dependent (PEP) phosphotransferase system (PTS) is a major carbohydrate uptake system in bacteria, which not only phosphorylates different carbohydrates during uptake but also plays a major a role in genetic regulation of metabolic activities (20, 29, 68, 142, 199, 276). In *S. mutans*, it has been shown that CelR has two PDR domains and phosphorylation at H226, H332 and H576 is required for the activation of CelR in the presence of cellobiose, while phosphorylation of H284 and H391 leads to inhibition of the CelR-dependent activation in the presence of glucose (275). Based on the CelR-II amino acid sequence, no such domains could be identified. Indeed, we could not find an effect of deletion of the *cel-II* PTS genes on the cellobiose and glucose-dependent regulation of P*celA* via CelR-II. Therefore, it remains to be determined how the signal of the carbohydrate source is transferred to CelR-II. In addition, no effect of CcpA on the regulation of *cel-II* operon

could be detected, in accordance with previous studies (43). Thus, the *celR-II* locus seems to be an autonomous unit concerning its transcriptional regulation.

As reported in our previous study the *cel* locus and the CelR regulatory site of *S. pneumoniae* were found highly conserved in other streptococci, although not all pneumococcal strains do contain this operon (168). Blast searches revealed a high conservation of the *cel-II* operon in all the strains of *S. pneumoniae* that are available on the KEGG website. However, no similar *cel-II* operon organization is found in any other *Streptococcus sp*. Thus, the *cel-II* operon seems to be specific for *S. pneumoniae*. Blast search with CelR-II revealed high homology of this protein with BgcR in *Escherichia coli*, which is a positive transcriptional regulator of a *bgc* operon (177). The *bgc* operon, which consist of 5 genes encoding PTS subunits IIB (BgcE), IIA (BgcF), IIC (BgcI) and phospho- β-glucosidase (BgcA) is found to be involved in the utilization of cellobiose and other β-glucosides (arbutin and salicin) at low temperature (177). The high homology of the *bgc* operon with the *cel-II* operon might indicate the involvement of the *cel-II* operon in the utilization of other β -glucosides like arbutin and salicin. Deletion of the genes encoding the PTS components (*celBCD-II*) of the *cel-II* operon has no effect on the growth compared to the wild-type in the presence of cellobiose as the carbon source, whereas deletion of *celR* (impairing expression of the *cel* locus) does. Thus, the *cel* locus is primarily important for cellobiose utilization in *S. pneumoniae* D39 and the exact role of the *cel-II* operon in the metabolism of β-glucosides remains to be determined.

To be able to identify more putative targets of CelR-II, we searched for the CelR-II operator site in the entire genome of *S. pneumoniae* D39 strain. This revealed that the operator site is only found in the promoter region of the *celA-II* gene. This suggests that the *cel-II* operon is the only direct target of CelR-II. However, deletion of *celR-II* also led to upregulation of a maltose utilization gene cluster. It is already known that the *mal* gene cluster is regulated by *malR*, a maltose-dependent transcriptional repressor (179, 180). Therefore, the upregulation of the *mal* gene cluster is probably an indirect effect of the high upregulation of the *cel-II* operon. We are currently also investigating the regulation mechanism of the *mal* gene cluster.

Acknowledgements

SS is supported by the HEC-NUFFIC overseas scholarship program by HEC (Higher Education commission) Pakistan and TGK is supported by the EU-funded Pneumopath project HEALTH-F3-2009-222983.

Chapter 6

Characterization of the ROK family transcriptional regulator RokA of *Streptococcus pneumoniae* D39

Sulman Shafeeq, Tomas G. Kloosterman, Vijayanand Rajendran and Oscar P. Kuipers

(*Submitted*)

Abstract

The Gram-positive human pathogen *Streptococcus pneumoniae* possesses an unusually high number of gene clusters specific for carbohydrate utilization. This provides it with the ability to use a wide array of sugars, which may aid during infection and survival in different environmental conditions present in the host. In this study, the regulatory mechanism of transcription of a gene cluster, *SPD0424-28*, putatively encoding a cellobiose/lactose-specific PTS, is investigated. We demonstrate that this gene cluster is transcribed as one transcriptional unit directed by the promoter of the *SPD0424* gene. Upstream of *SPD0424* a gene was identified encoding a ROK-family transcriptional regulator (RokA: SPD0423). DNA microarray and transcriptional reporter analyses with a *rokA* mutant revealed that RokA acts as a transcriptional repressor of the *SPD0424-28* operon. Furthermore, we identified a 25-bp AT-rich DNA operator site (5'-TATATTTAATTTATAAAAAATAAAA-3') in the promoter region of *SPD0424*, which was validated by promoter truncation studies, DNAseI footprinting and electrophoretic-mobility shift assays (EMSAs). We tested a large range of different sugars for their effect on the expression of the *SPD0424-28* operon, but only very moderate variation in expression was observed in the conditions applied. Therefore, a co-factor for RokA-mediated transcriptional control could not be identified.

Introduction

Streptococcus pneumoniae is an opportunistic Gram-positive human pathogen that resides as a commensal in the nasopharynx (31). During favorable conditions, it has the ability to spread to different parts of the human body, where it can cause serious infections like otitis media, pneumonia, meningitis and bacteremia (114). To be able to survive in the host, it must adapt to the different environmental conditions at diverse anatomical sites and tune itself to a fluctuating availability of nutrients in the host, an important one of which is the carbon/energy source (31, 114). This requires regulatory mechanisms, ensuring that carbon utilization systems necessary for growth under certain conditions are properly expressed.

Glucose is the preferred carbon source for *S. pneumoniae*, as it leads to the shortest doubling time (29, 43, 103). However, *S. pneumoniae* has the ability to grow on various other carbon sources as well, when provided as a sole energy source in the medium (29, 43, 103, 159, 168, 223, 250, 250). Compared to human blood (3.57-6.06 mM) glucose levels at other common sites of pneumococcal infection are quite low (saliva 0.02-0.4 mM (82), nasal secretions <1.0 mM (269) and lower airway secretions <0.5mM (60)), meaning that

alternative carbon sources are engaged at these sites to maintain the fitness of *S. pneumoniae* during infection (228). Indeed, more than 30% of the transporters in *S. pneumoniae* are predicted to be involved in carbohydrate transport (29). These transport systems include phosphoenolpyruvate (PEP)-dependent phosphotransferase systems (PTS), ATP-binding cassette (ABC) and ion gradient-driven transporters and enable *S. pneumoniae* to utilize different types of carbohydrates in the host (29, 43, 49, 103, 104, 159, 168, 180, 223).

It is important for the cell to strictly control the expression of carbohydrate utilization systems, in order to channel the cell's energy towards the metabolism of only the carbon source that is available at a certain moment. In *S. pneumoniae* and other bacteria, the ability to use the preferred sugar(s) before the non-preferred sugar(s) depends on a regulatory process called carbon-catabolite repression (CCR) (43, 80, 103). CCR brings about silencing/repression of genes that are specific for utilization of non-preferred sugars, until the cell has completely consumed the preferred carbon source (43, 103). In addition to CCR, the regulation of some sugar-uptake systems is governed by dedicated regulatory factors, which guarantee expression of these sugar uptake systems at the right moment (104, 168, 179, 223, 250).

One class of transcriptional regulators involved in carbohydrate-dependent transcriptional control is a member of the ROK (Repressor, Open reading frame, and Kinase) protein family, which also comprises sugar kinases and many functionally uncharacterized proteins (245). The transcriptional regulation mediated by ROK-family repressors and their role in carbohydrate utilization has been well elucidated in many bacteria (64, 71, 120, 132, 194-196). In *Escherichia coli* the ROK-family transcription factor Mic is a repressor protein of genes and operons (*ptsG, ptsHI* and *manXYZ)* involved in glucose utilization (120) and maltose metabolism (64). Other examples are NagC, a repressor of the N-acetylglucosamine-(GlcNAc) utilization operon (*nagE-nagBACD*) (196), the *chb* operon involved in the utilization of chitobiose (195), and an activator of the *glmUS* operon involved in synthesis of UDP-GlcNAc, the essential precursor for cell wall components (194). In *Bacillus subtilis*, a ROK-family repressor, XylR, was characterized as a repressor of xylose utilization genes (132).

Recently, a ROK-family transcriptional repressor has been demonstrated to be involved in the repression of a cellobiose/lactose PTS operon and sulfatase-encoding genes in *S. pneumoniae*. Moreover, it was found that these genes contribute to the pathogenesis of *S. pneumoniae* strain WCH206 (164). This specific ROK-family protein is encoded on a genomic island, which is however absent in many pneumococcal strains, including D39.

Despite the absence of this genomic island, the genome of *S. pneumonia*e D39 encodes four ROK-family proteins, one of which codes for a ROK-family transcriptional repressor protein, RokA, which is not conserved in all pneumococcal strains. In this study, we report the functional characterization of RokA in strain D39 and show that it is directly involved in the regulation of a putative cellobiose/lactose PTS operon.

Material and Methods

Table 1: List of strains and plasmids used in this study.

Strain/plasmid	Description	Source
S. pneumoniae		
D39	Serotype 2 strain, *cps 2*	Laboratory of P. Hermans
SS500	D39 Δ*rokA*; $Spec^R$	This study
SS501	D39 Δ*bgaA*::P*spd0424-lacZ*; Tet^R	This study
SS502	D39 Δ*bgaA*::P*spd0426-lacZ*; Tet^R	This study
SS503	SS500 Δ*bgaA*::P*spd0424-lacZ*; Tet^R	This study
SS504	SS500 Δ*bgaA*::P*spd0426-lacZ*; Tet^R	This study
SS505	D39 Δ*bgaA*::P*spd0424-P1-lacZ*; Tet^R	This study
SS506	D39 Δ*bgaA*::P*spd0424-P2-lacZ*; Tet^R	This study
SS507	D39 Δ*bgaA*::P*spd0424-P3-lacZ*; Tet^R	This study
SS508	D39 Δ*bgaA*::P*spd0424-P4-lacZ*; Tet^R	This study
SS509	SS500 Δ*bgaA*::P*spd0424-P1-lacZ*; Tet^R	This study
SS510	SS500 Δ*bgaA*::P*spd0424-P2-lacZ*; Tet^R	This study
SS511	SS500 Δ*bgaA*::P*spd0424-P3-lacZ*; Tet^R	This study
SS512	SS500 Δ*bgaA*::P*spd0424-P4-lacZ*; Tet^R	This study
L. lactis		
NZ9000	MG1363 Δ*pepN::nisRK*	(134)
SS513	NZ9000:: pSS507	This Study
E. coli		
EC1000	Km^R; MC1000 derivative carrying a single copy of the pWV1 *repA* gene in *glgB*	Laboratory collection
Plasmids		
pPP2	Amp^R Tet^R; promoter-less *lacZ*. For replacement of *bgaA* with promoter *lacZ* fusion. Derivative of pTP1	(88)
pNZ8048	Cm^R: Nisin inducible P*nisA*	(61)
pSS501	pPP2 P*spd0424-lacZ*	This study
pSS502	pPP2 P*spd0426-lacZ*	This study
pSS503	pPP2 P*spd0424-P1-lacZ*	This study
pSS504	pPP2 P*spd0424-P2-lacZ*	This study
pSS505	pPP2 P*spd0424-P3-lacZ*	This study
pSS506	pPP2 P*spd0424-P4-lacZ*	This study
pSS507	pNZ8048 carrying Strep-tagged RokA in downstream of P*nisA*	This study

General procedures

Strains and plasmids used in this study were stored in 10% (v/v) glycerol at -80 °C and are listed in Table 1. *S. pneumonia*e (136) was grown in M17 broth (242) supplemented with 0.5% (w/v) glucose, on blood agar plates supplemented with 1% (v/v) defibrinated sheep blood in micro-aerophilic conditions at 37 °C. LB broth was used to grow *Escherichia coli* in a shaker incubator at 37 °C. *Lactococcus lactis* strain NZ9000 was used for overexpression of the RokA protein using the nisin-inducible system and was grown at 30 °C in GM17 broth as described before (134). When appropriate, media were supplemented with the following concentrations of antibiotics: spectinomycin: 150 μg ml^{-1}, tetracycline: 2.5 μg ml^{-1} for *S. pneumoniae*; ampicillin: 100 μg ml^{-1} for *E. coli* and chloramphenicol: 4 μg ml^{-1} for *L. lactis*. Nisin (sigma) was used to induce nisin-dependent overexpression at a concentration of 10 ng ml^{-1} in *L. lactis* (126, 133, 134). Chromosomal DNA of *S. pneumoniae* wild-type strain D39 was used as a template for PCR amplification (15, 136) and all DNA manipulations were done as described (125). Primers used in this study are based on the sequence of the D39 genome (136) and listed in Table 2.

Table 2: List of primers used in this study.

Name	Nucleotide sequence (5' to 3')[a]	Restriction site
PSPD0424-F	CATGGAATTCGTTGCTGGATTTCGATATAG	EcoRI
PSPD0424-R	CATGGGATCCGACTTTTGTTTCTAAAAACGC	BamHI
PSPD0426-F	CATGGAATTCGGATGTTTATCGTGTAATTG	EcoRI
PSPD0426-R	CATGGGATCCGACAATTGTAAAACCAAG	BamHI
PSPD0424-F1	CATGGAATTCAAGGCTGATTGTTTTGTAATC	EcoRI
PSPD0424-R1	CATGGGATCCTTCATCACCGGAACTCC	BamHI
PSPD0424-F2	CATGGAATTCTTAAACATAAGAAAAAATC	EcoRI
PSPD0424-R2	CATGGGATCCAATATACCACTAACAGAAAG	BamHI
PSPD1830-F	CGGGATCCCCGCTAGAAGCTGCTCCCCACC	EcoRI
PSPD1830-R	CGGAATTCCTTTTCACGAATCTCATTGT	BamHI
SPD0423-1	GAAGACTAGGAATGCCAAAACACTGAC	-
SPD0423-2	GCATGCGGCCGCCCTTAATATATTAGAAATCTATAGACCTG	NotI
SPD0423-3	TGCTAGGCGCGCCTGTAATCATTTGTTTCCTCC	AscI
SPD0423-4	CAACTCCTGCCGCTTCAAGACCT	-
Spec-F	GCATAGGCGCGCCCTAATCAAAATAGTGAGGAGGATAT	AscI
Spec-R	CGATTGCGGCCGCACTAAACGAAATAAACGCTAAAACG	NotI
SPD0423-nco	CGATCCATGGTAAAGGAGGAAACAAATGATTAC	NcoI
SPD0423- Cstrep-hin	GATCAAGCTTATTTTTCAAATTGTGGATGGCTCCAAGCGCT AGGTGAAATAATATCTTG	HindIII

[a]Restriction enzyme sites are underlined.

Construction of the *rokA* mutant

The *rokA* (*SPD0423*) mutant was made by means of allelic replacement with a spectinomycin-resistance cassette. In short, primers SPD0423-1/SPD0423-2 and SPD0423-3/SPD0423-4 were used to generate PCR products of the left- and right flanking regions of

rokA, which were, by means of ligation using AscI/NotI restriction sites, fused to a spectinomycin-resistance gene, PCR amplified with primers Spec-F and Spec-R from plasmid pORI38 (140). The resulting ligation mixture was transformed to the *S. pneumoniae* strain D39, yielding strain SS500, and the mutation was further verified by PCR.

Construction of transcriptional *lacZ*-fusions of the *SPD0424* and *SPD0426* promoters

Chromosomal transcriptional *lacZ*-fusions to the *SPD0424* and *SPD0426* promoters were constructed in the integration plasmid pPP2 (88) with the primer pairs PSPD0424-F/PSPD0424-R and PSPD0426-F/PSPD0426-R, respectively (Table 2), resulting in plasmids pSS501-02. These plasmids were further introduced into wild-type D39 and the *rokA* mutant (SS500), resulting in strains SS501-04. All plasmid constructs were checked by DNA sequencing

Subcloning of the *SPD0424* promoter

A scheme of P*SPD0424* subclones is shown in Fig. 2A. Transcriptional *lacZ*-fusions of subclones P1, P2, P3 and P4 were constructed in plasmid pPP2 (88) with the primer pairs PSPD0424-F1/PSPD0424-R1, PSPD0424-F2/PSPD0424-R1, PSPD0424-F1/PSPD0424-R2 and PSPD0424-F2/PSPD0424-R2, respectively, resulting in plasmids pSS503-06. These plasmids were further transformed into *S. pneumoniae* wild-type D39 and the *rokA* mutant (SS500), giving strains SS505-12. All plasmid constructs were checked by DNA sequencing.

β-galactosidase assays

For β-galactosidase assays cells were grown in triplicate in M17 broth at 37 °C in the presence of 0.5% (w/v) of one of the following sugars: glucose, fructose, mannose, arabinose, mannitol, cellobiose, dextrose, maltose, sucrose, raffinose, lactose, trehalose, N-acetyl glucosamine or N-acetyl galactosamine. Cells were harvested in the mid-exponential phase of growth and specific β-galactosidase activity was measured as described (126).

Overexpression and purification of RokA

To overexpress C-terminally Strep-tagged RokA, primers SPD0423-nco/SPD0423-Cstrep-hin were used for the PCR amplification of the *rokA* gene. The *rokA* PCR product was restricted with NcoI/HindIII and cloned into the NcoI/HindII sites of pNZ8048 (62), resulting in plasmid pSS507. Overexpression of C-terminally Strep-tagged RokA was achieved in strain *L. lactis* NZ9000 as described before (134, 224).

Electrophoretic mobility shift assays (EMSAs) and DNAaseI footprinting analysis

Electrophoretic mobility shift assays (EMSAs) were performed as described before (224). PCR products of P*SPD0424*, P*SPD0426*, P*SPD1830* and truncated fragments of P*SPD0424* (P1, P2, P3 and P4) were amplified with primer pairs PSPD0424-F/PSPD0424-R, PSPD0426-F/PSPD0426-R, PSPD0424-F1/PSPD0424-R1 (P1), PSPD0424-F2/PSPD0424-R1 (P2), PSPD0424-F1/PSPD0424-R2 (P3) and PSPD0424-F2/PSPD0424-R2 (P3), respectively and were labeled with [γ-^{33}P]ATP. A PCR fragment of P*SPD1830* that was amplified with primer pair PSPD1830-F/PSPD01830-R was used as a negative control. 3000 cpm per reaction of [γ-^{33}P]ATP-labeled PCR products were used to perform EMSAs with increasing concentration of Strep-tagged RokA.

Experimental procedures for DNAaseI footprinting were performed essentially as described before (65, 128). For each reaction, 150,000 cpm [γ-^{33}P]ATP-labeled PCR product of P*SPD0424* that was PCR amplified by either [γ-^{33}P]ATP-labeled primer SPD0424-F (forward strand) or SPD0424-R (reverse strand) in combination with the unlabeled primer, was used.

DNA microarray experiments and data analysis

For microarray analysis of *S. pneumoniae* wild-type and its isogenic *rokA* mutant (SS500), cells were grown in 3 biological replicates in GM17 and harvested at an OD600 of approximately 0.3. All other procedures regarding microarrays and data analysis were done as described (225, 254). The DNA microarray data have been submitted to GEO.

Results

In silico* analysis of ROK-family protein members in *S. pneumoniae

The ROK-family of protein consists of transcriptional repressors, kinases and uncharacterized proteins (54, 245). The ROK-family repressor proteins are generally about 400 amino acids (aa) in length while the kinases are about 100 aa shorter due to the absence of a helix-turn-helix (HTH) DNA binding motif (54). Blast search for ROK-family proteins in the genome of *S. pneumoniae* D39 strain revealed the presence of four ROK-family proteins in D39. Here, these proteins are named RokA (SPD0423), RokB (SPD0580), RokC (SPD1488), and RokD (SPD1970) (Table 3). The *rokA* gene, which is present upstream of a PTS gene cluster (*SPD0424-28*) (Fig. 1A), encodes a 407 amino acids (aa) protein with an N-terminal HTH (Helix-Turn-Helix) DNA binding motif. The second ROK protein, RokB, is a

putative glucosekinase and is about 90 aa shorter than RokA, due to the absence of an N-Terminal HTH DNA binding motif. RokC is located in close proximity to a gene (*SPD1489*) that is involved in sialic acid utilization and metabolism (160). The fourth ROK protein, RokD, is encoded in an operon with three other hypothetical genes. The function of this operon is unknown. However, this operon was highly upregulated in the presence of cellobiose or in the absence of glucose (Chapter 5). Thus, RokA seems to be the only ROK protein in *S. pneumoniae* that has an N-Terminal HTH DNA binding motif. It is therefore to be expected that it can act as a transcriptional repressor, possibly involved in carbohydrate metabolism. Since the fitness of *S. pneumoniae* likely requires proper ways to ensure optimal expression of its sugar metabolic genes, we started to investigate the role of RokA and the gene cluster with which we hypothesized it to be associated.

Table 3: List of ROK family proteins present in the genome of *S. pneumoniae* D39 wild-type strain.

D39 locus Tag	TIGR4 locus Tag	Description	Residues (aa)
SPD0423	*SP0473*	ROK family protein (RokA)	407
SPD0580	*SP0668*	ROK family protein (RokB)	319
SPD1488	*SP1675*	ROK family protein (RokC)	294
SPD1970	*SP2142*	ROK family protein (RokD)	289

Organization of gene cluster *SPD0424-28*

The organization of the gene cluster lying adjacent to *rokA*, *SPD0424-28*, is shown in Fig. 1A. It consists of five genes and is most likely organized in one transcriptional unit. These genes encode the following proteins: *SPD0424* encodes a cellobiose/lactose specific PTS IIC component, *SPD0425* encodes a hypothetical protein, *SPD0426* encodes a cellobiose/lactose specific PTS IIC component, *SPD0427* encodes an enzyme 6-phospho-β-galactosidase and *SPD0428* is predicted to encode cellobiose/lactose-specific PTS IIBC components.

To investigate whether this gene cluster transcribes into one transcriptional unit, transcriptional *lacZ*-fusions of both intergenic regions (P*SPD0424* and P*SPD0426*) were constructed in pPP2 plasmid and introduced into D39 wild-type. β-galactosidase assays showed that from the strains containing these transcriptional *lacZ*-fusions only the one with the P*SPD0424-lacZ* construct was active, while no expression of P*SPD0426-lacZ* was observed in GM17 medium (Fig. 1B). Indeed, no clear promoter sequences could be identified in P*SPD0426-lacZ*, whereas we could find obvious -35 and -10 sites in P*SPD0424* (Fig. 1A). These data suggest that most likely this gene cluster (*SPD0424-28*) is transcribed as one transcriptional unit. Thus, from hereon, it will be denoted as an operon.

RokA represses the expression of the *SPD0424-28* operon

To explore whether RokA is involved in the regulation of the *SPD0424-28* operon, the *rokA* gene was replaced by a spectinomycin-resistance marker. P*SPD0424-lacZ* and P*SPD0426-lacZ* transcriptional fusions were transformed into the *rokA* mutant strain and β-galactosidase assays were performed with the strains containing these transcriptional *lacZ*-fusions grown in GM17 (0.5% Glucose + M17) medium. Deletion of *rokA* led to high expression of P*SPD424-lacZ,* which suggests a role of RokA as a transcriptional repressor of the *SPD0424-28* operon (Fig. 1B). No expression was observed with the P*SPD0426-lacZ* in the *rokA* mutant strain. This confirms that SPD*0424-28* constitutes one transcriptional unit arising from P*SPD0424* (Fig. 1B).

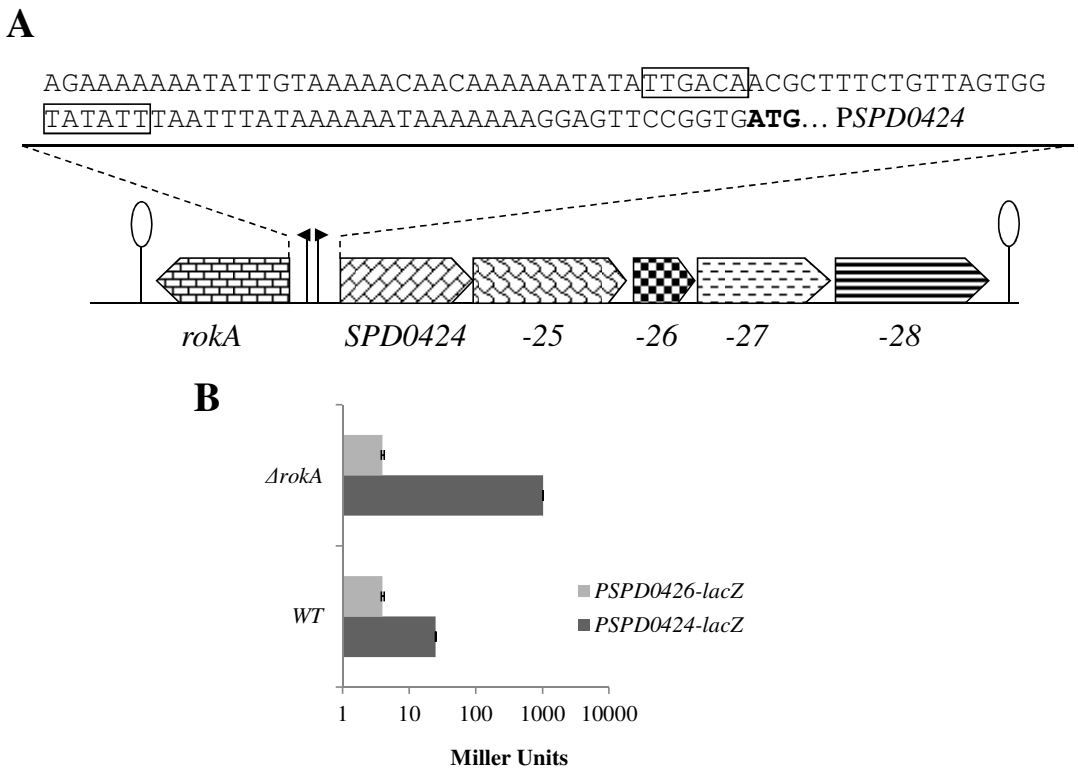


Figure 1: A) Organization of the *rokA* gene and its upstream operon (*SPD0424-28*). Flags indicate the presence of putative promoters while ovals indicate the presence of putative terminators. Sequence of the P*SPD0424* is indicated. Translation initiation codon (ATG) is italic and the predicted -35 and -10 core promoter sequences are in rectangular, **B)** Specific β-galactosidase activity (Miller Units) of D39 wild-type and *rokA* mutant strains containing the P*SPD0424-lacZ* and P*SPD0426-lacZ* transcriptional fusions grown in GM17 medium.

DNA microarray analysis with the Δ*rokA* mutant strain

To investigate the effect of *rokA* deletion on the transcriptome of *S. pneumoniae* D39, the wild-type was compared to its isogenic *rokA* deletion strain grown in GM17 (0.5% Glucose + M17) medium. Table 4 summarizes the transcriptome changes induced upon

deletion of *rokA*. The *rokA* deletion seems to have a very specific effect on the transcriptome of *S. pneumoniae* D39, since after applying the criteria of ≥ 3.0-fold difference as the threshold change and a *P* value < 0.001, the *SPD0424-28* operon was the only cluster of genes significantly upregulated in the *rokA* deletion strain. This confirms the function of RokA as a repressor of the *SPD0424-28* operon and suggests that the *SPD0424-28* operon is the only target of RokA.

Table 4: Summary of transcriptome comparison of *S. pneumoniae* wild-type strain D39 with the isogenic *rokA* mutant (SS500) grown in GM17 medium. [a]Ratios >2.0 or <-2.0 (D39 *ΔcelR-II* compared to D39 wild-type.

D39 locus tag	Function	[a]Ratio
SPD0423	ROK family transcriptional repressor RokA	-23.6
SPD0424	PTS system, cellobiose-specific IIC component	57.2
SPD0425	Hypothetical protein	84.8
SPD0426	LacF-1; PTS system, lactose-specific IIA component	33.6
SPD0427	LacG-1; 6-phospho-β-galactosidase	112.8
SPD0428	LacE-1; PTS system, lactose-specific IIBC components	48.6

Identification of a RokA operator site in P*SPD0424*

To find the putative operator site for RokA, a series of truncations from the 5'- and 3' end of P*SPD0424* (P1, P2, P3 and P4) was constructed (Fig. 2A). These truncated promoter fragments were transcriptionally fused to *lacZ* and transformed to D39 wild-type and the *rokA* mutant. β-galactosidase measurements revealed that deletion of the region from base-pair -240 to -188 (P1) and from base-pair -188 to -139 (P2) relative to the start codon of *SPD0424* gene, did not show any significant increase in expression of P1 and P2 compared to P0 (full promoter of P*SPD0424*) (Fig. 2B). However, deletion of the -34 to +1 (P3 and P4) region, relative to the start codon of *SPD0424* gene, led to derepression of P*SPD0424* expression in D39 wild-type. In the *rokA* mutant, derepression of all truncations (P1, P2, P3 and P4) was observed compared to the wild-type (Fig. 2B). This suggests the presence of a RokA operator site in the -34/+1 region that mediates RokA-dependent repression of the *SPD0424-28* operon.

Verification of the RokA binding site by means of the EMSAs and DNAseI footprinting

To study the direct interaction of RokA with its target promoters, we performed EMSA experiments with purified strep-tagged RokA. These assays indicated that RokA binds specifically to P*SPD0424*, but not to P*SPD0426* (Fig. 3), which is consistent with the above-mentioned results (Fig. 3). P*SPD1829* was used as a negative control. To find out whether the identified region (-34 to +1) mediates the RokA-dependent regulation of P*SPD0424*, we performed EMSA experiments with truncated fragments of P*SPD0424* (see Fig. 2A for a

schematic overview of the truncations). When the -34 to +1 region of P*SPD0424* was deleted (P3 and P6), binding of RokA was completely abolished (Fig. 3) even at higher concentrations of RokA-Strep. However, RokA-Strep did bind to promoter fragments of P*SPD0424* that do contain the -34 to +1 region (*i.e.* P1, P2 and P5; Fig. 3). Thus, these RokA-P*SPD0424* interaction assays demonstrate that RokA is directly involved in the control of P*SPD0424*, and in addition suggest the occurrence of an operator site for RokA in the -34 to +1 region relative to the start codon of *SPD0424*, which is fully consistent with the transcriptional reporter analysis described above.

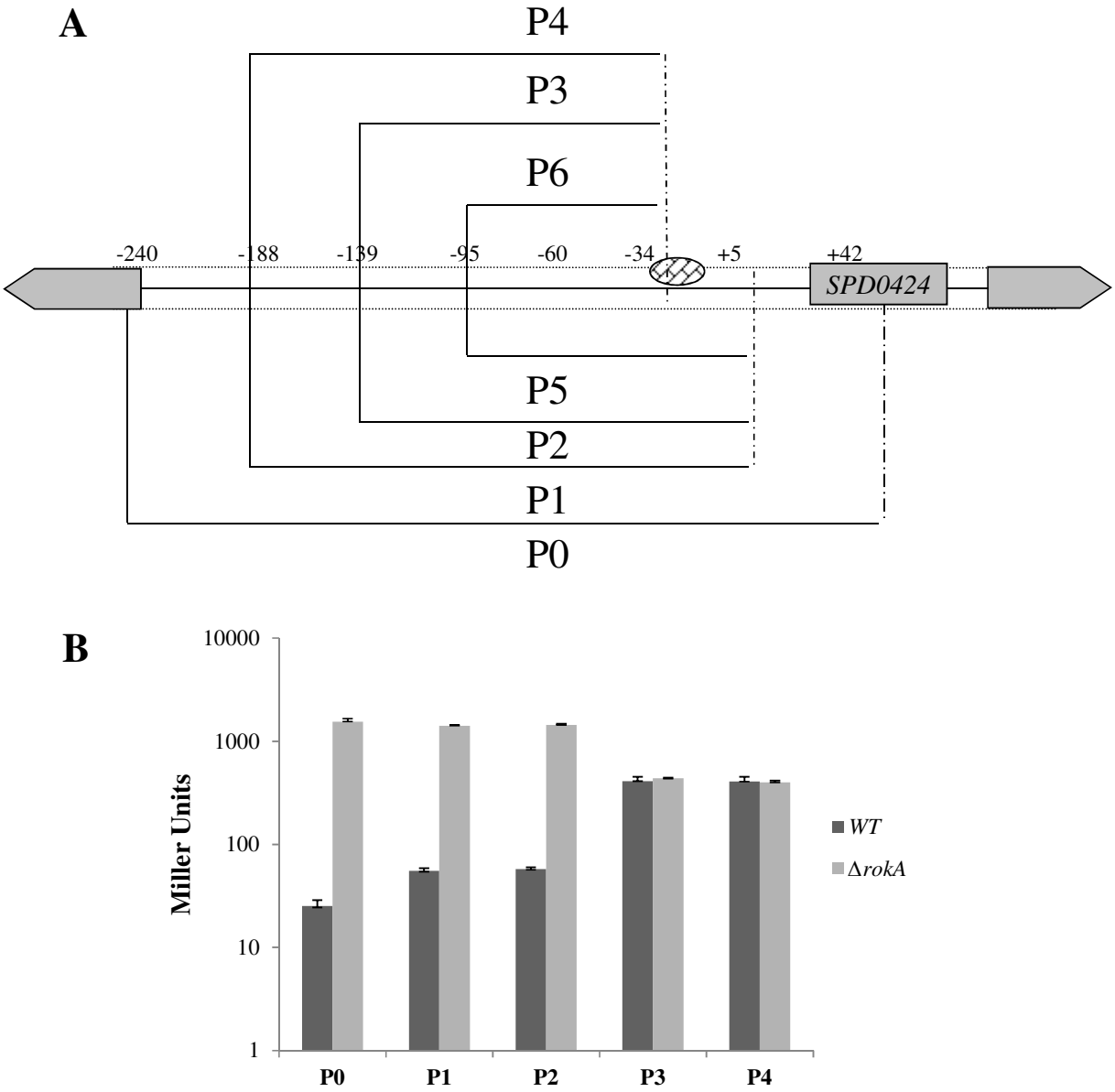


Figure 2: Subcloning of P*SPD0424*. **A)** A schematic drawing of P*SPD0424* truncations is shown from both 5'- and 3' ends. **B)** Expression (Specific β-galactosidase activity in Miller Units) of the truncated P*SPD0424* promoter fragments in *S. pneumoniae* wild-type D39 and *rokA* mutant strains grown in GM17 medium, as measured by transcriptional fusions to *lacZ*.

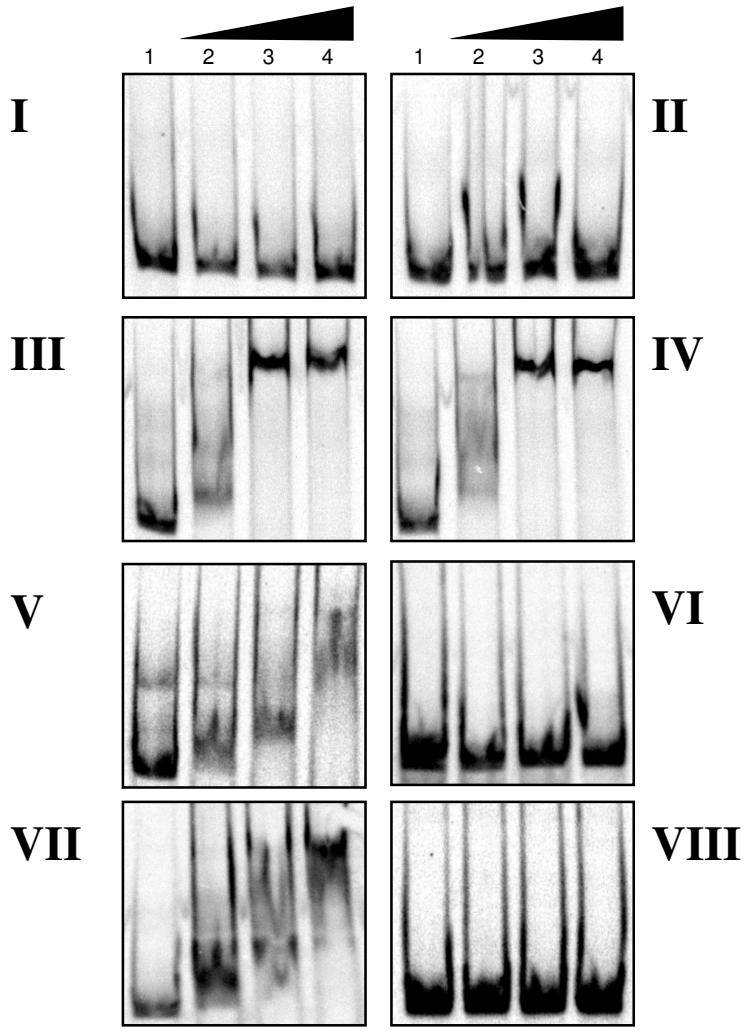


Figure 3: *In vitro* interaction of RokA-Strep with the promoter regions of *SPD1829* (**I**), *SPD0426* (**II**), and the full-length and truncated *SPD0424* promoter (*P0* (**III**), *P1* (**IV**), *P2* (**V**), *P3* (**VI**), *P5* (**VII**), and *P6* (**VIII**)). Purified RokA-Strep was added from 0 to 500 nM in lanes 1 to 4, respectively. The triangular bar above lanes 2 to 4 indicates the increasing concentration of RokA-Strep. Arrows indicate the position of the shifted probes. The presence of weaker bands for some of the DNA fragments that run higher than the free probe in the gels is a phenomenon that has also been seen by others in similar experiments. These bands may represent unspecific PCR products or single-stranded DNA (4, 66).

We next wanted to identify more exactly the RokA recognition sequence. Therefore, a DNAseI footprint experiment was performed. For this purpose, the forward strand of P*SPD0424* was radio labeled with [γ-^{33}P]ATP and incubated with increasing concentrations of RokA-Strep. A very clear region of protection was observed that lies within the area spanning bp -34 to +1 of the *SPD0424* promoter (Fig. 4). When the DNAseI footprinting experiment was performed with the labeled complementary DNA strand, a similar protection pattern was seen (data not shown). Further analysis of this protected region revealed a 25-bp AT rich possible operator site (5'-TATATTTAATTTATAAAAAATAAAA-3') (Fig. 4), with

resemblance to the predicted operator site of XylR (5'-ACTTATTAAANNNNNTTTAAAAAGT-3') in Firmicutes (84), which we hypothesize to function as the RokA operator.

This 25-bp operator sequence was subsequently used to search the entire genome of *S. pneumoniae* D39 strain with the Genome 2D software (16). This search revealed that the RokA operator site is exclusively present in P*SPD0424* even with 5 allocated mismatches, suggesting that the *SPD0424-28* operon is the only direct target of RokA in *S. pneumoniae*.

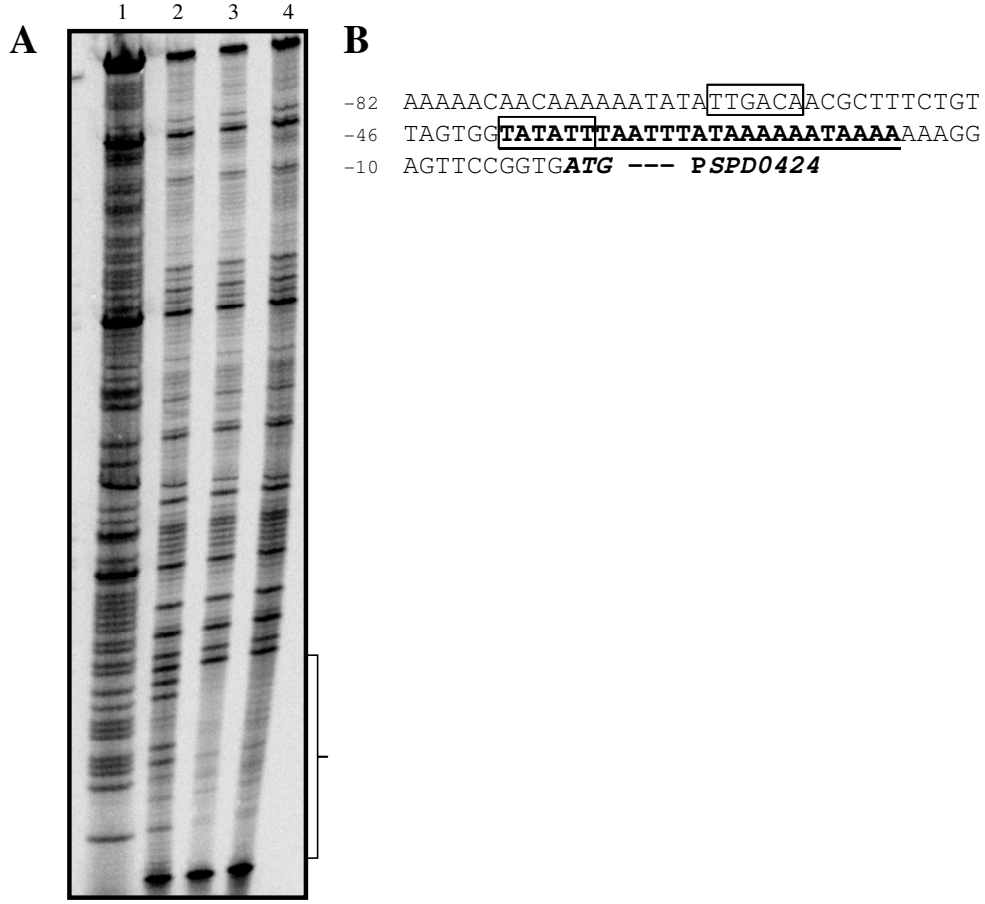


Figure 4: DNAase I footprinting analysis of RokA binding to P*SPD0424* **A)** Radioactively labeled probe comprising forward stand of P*SPD0424* was DNAseI treated alone (lane 2) or in the presence of 100 nM (lane 3) or 200 nM (lane 4) of RokA-Strep. Footprints are flanked on the left by Maxam and Gilbert A+G sequence ladders (lane 1). The protected region is marked with a bar in the right. **B)** Sequence of P*SPD0424*. The proposed RokA operator site is mentioned in bold and underlined. The predicted -35 and -10 core promoter sequences are in rectangular.

RokA seems to mediate its repression by a 25-bp AT-rich operator site present in P*SPD0424*. Blast searches with RokA revealed that RokA homolog are present in various streptococci. Therefore, it was investigated whether the RokA operator site inferred from our experiments is conserved in these streptococci as well. Interestingly, a possible RokA

operator site was found in streptococci that contain a similar composition of the *SPD0424-28* operon as in *S. pneumoniae* (Fig. 5). The predicted RokA operator site from these streptococci was aligned with that of *S. pneumoniae* and a 25-bp putative consensus sequence was generated (TATATTTWATTTATAAAAWAAAAAW) (Fig. 5). The presence of a similar operon composition and operator site in these streptococci suggests a regulatory role of RokA that is similar to what is found in this study.

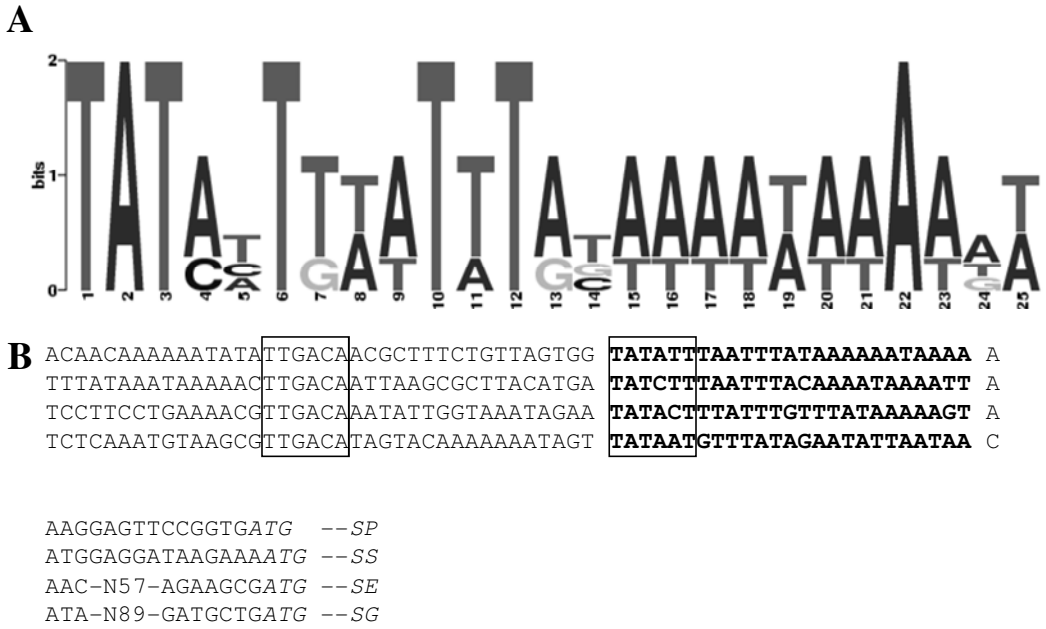


Figure 5: Identification of RokA operator site in different streptococci. **A)** Weight matrix of the identified RokA operator site in different streptococci. **B)** Position of operator site in P*SPD0424* of different streptococci. Where the predicted RokA operator site is bold, the predicted -35 and -10 core promoter sequences are in rectangular and translation initiation codon is italic. SP= *S. pneumoniae*, SS= *S. suis*, SE= *S. epui* and SG= *S gallolyticus*.

Search for a co-factor for RokA-mediated regulation

Based on the fact that RokA regulates an operon containing genes that encode a PTS system, we hypothesized that a specific carbohydrate compound could lead to relieve of RokA-mediated repression of P*SPD0424*. To search for a co-factor for RokA, β-galactosidase assays were performed with P*SPD0424-lacZ* in M17 medium with addition of 0.5% (w/v) of different carbon sources (arabinose, cellobiose, dextrose, galactose, fructose, lactose, glucose, maltose, mannose, mannitol, N-acetyl glucosamine (GlcNAc), N-acetyl-galactosamine (GalNAc), raffinose, sorbitol, sucrose, trehalose, xylose and mucin) (Fig. 6). Interestingly, none of the tested sugars was able to strongly induce the expression of P*SPD0424-lacZ* (Fig. 6), although some of the sugars caused a slightly increased expression of P*SPD0424-lacZ* (like raffinose), while others (e.g. glucose) led to lower expression as compared to M17

medium without addition of any carbon source (Fig. 6). It might be that the complex medium M17 contains a certain carbohydrate that prevents is a co-repressor of RokA, thereby preventing expression of P*SPD0424-lacZ*. Therefore, we checked the expression of P*SPD0424-lacZ* in chemically defined medium in the presence of different sugars. Interestingly, no difference in expression of P*SPD0424-lacZ* was observed as compared to M17 medium in the presence of different sugars (data not shown). It might be that a derivative or analogue of the sugars that were tested is the 'natural' ligand for RokA.

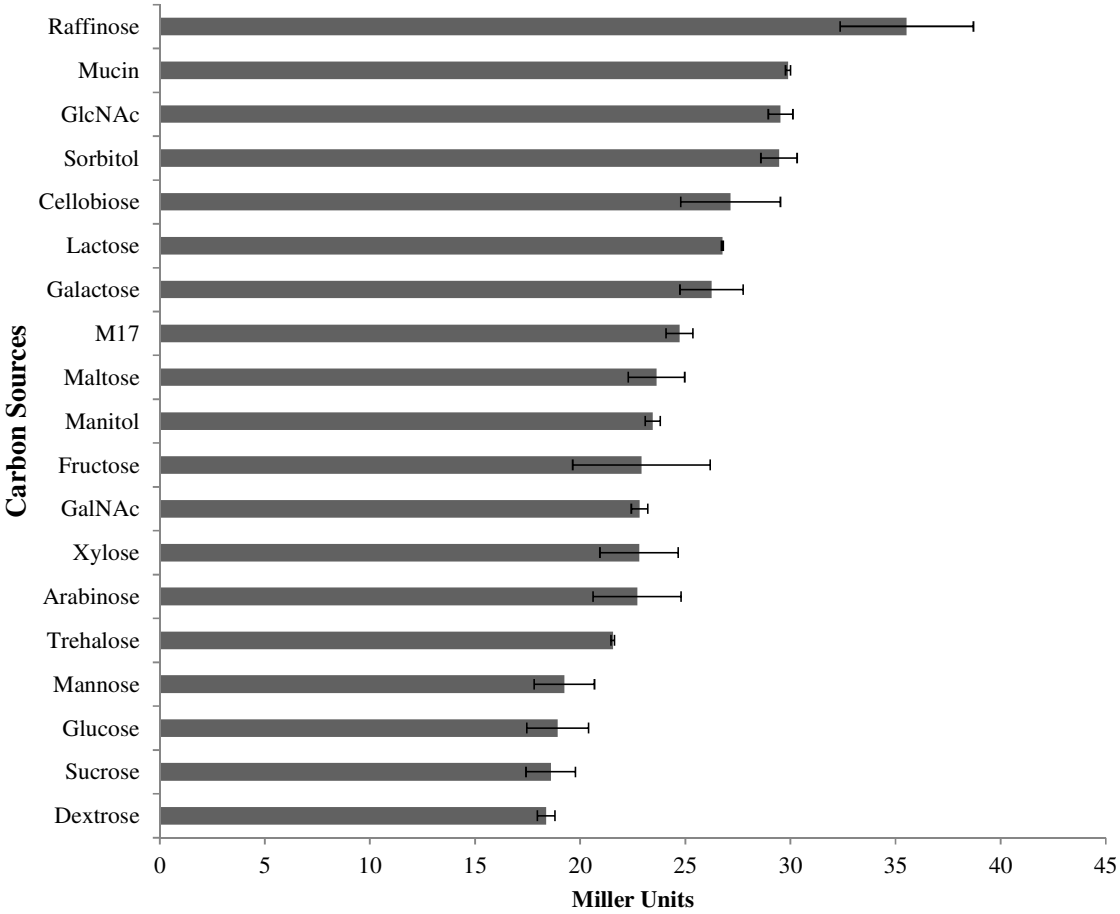


Figure 6: Specific β-galactosidase activity (Miller Units) of D39 wild-type containing the P*SPD0424-lacZ* grown in M17 medium with different added carbon sources (0.5% w/v).

Discussion

Many previous studies have indicated the ability of *S. pneumoniae* to utilize and respond to different sources of carbon (29, 43, 49, 103, 104, 159, 168, 180, 223). However, the regulatory mechanism of many carbohydrate utilization PTS systems has not been characterized yet in *S. pneumoniae* strain D39. The ROK-family is a group of proteins that comprises transcriptional repressors, sugar kinases and uncharacterized open reading frames (ORFs) (54, 245). The role of ROK-family proteins in carbohydrate utilization has been

already investigated in various bacteria including *B. subtilis*, *E. coli* and *S. pneumoniae* (64, 71, 120, 132, 194-196). In this study, the role of the ROK-family transcriptional repressor RokA was investigated, which we hypothesized to be involved in carbohydrate utilization in *S. pneumoniae* strain D39. Here, we showed that RokA acts as a transcriptional repressor of the *SPD0424-28* operon and this repression is mediated by direct binding of RokA to a 25-bp AT-rich DNA operator sequence present in P*SPD0424*. Despite the fact that the *SPD0424-28* operon has high sequence homology with other streptococcal lactose/cellobiose specific PTSs, we were not able to identify a possible co-factor that can cause derepression of the *SPD0424-28* operon.

In *S. pneumoniae*, regulation of carbohydrate utilization is not only important for proper metabolic functioning of the cell, but is also important for the pathogenesis. Like in other low-GC Gram-positive bacteria, CcpA plays a major role in regulation of carbohydrate utilization and virulence genes in *S. pneumoniae* (43). However, there are some carbohydrate utilization systems known to be regulated in a CcpA-independent fashion (104, 179, 223). Also the *SPD0424-28* operon seems not to be regulated by CcpA, based on DNA microarray analyses with a *ccpA* mutant in the presence of glucose and galactose (43). However, P*SPD0424* does contain a putative *cre* box (5'- CAGAAAGCGTTGTCAA -3') (181). It might be that this *cre* site is not functional or that the CcpA effect is only evident when RokA is not repressing the operon.

The main differences between ROK-kinases and -repressor proteins are the presence of a conserved N-terminal ATP-binding motif DxGxT and the absence of an N-terminal HTH DNA-binding motif in the kinases (54). Protein sequence alignment of RokA, B, C and D of *S. pneumoniae* strain D39 revealed that an N-terminal ATP-binding motif DxGxT is present in RokB, C and D, while the N-terminal HTH DNA-binding motif is only present in RokA. The presence of an ATP-binding motif DxGxT might suggest a kinase function of RokB, C and D proteins. Therefore, further investigation of these proteins might help to understand their role in carbohydrate utilization. Moreover, ROK-family proteins also have a conserved metal binding site (CxCGxxGCx(E/D)) that coordinates a single atom of zinc (54). Protein sequence alignment of RokA, B, C and D revealed that this metal binding site is only conserved in RokA and B while absent in RokC and D. This suggests an important role of zinc in the structure and proper functioning of ROK proteins.

The number of PTS system varies from 15-20 between different pneumococcal strains and many of the PTS systems are not conserved in all the strains of *S. pneumoniae* (29). Blast searches revealed that the *SPD0424-28* operon is not conserved in all the pneumococcal

strains (D39, R6, TIGR4, G54, Hungary 19F-6, 70585, P1031, Taiwan 19F, TCH8431 and 670-6B) available on the KEGG website (www.genome.jp/kegg/) as it is absent in strains ATCC700669, JJA and CGSP14. However, recently the role of a ROK-family transcriptional regulator and its upstream operons (a putative cellobiose/lactose PTS operon and sulfatase operon), present on a genomic island, in the pathogenesis of *S. pneumoniae* strain WCH206 was studied (164). This genomic island is present in strains ATCC700669, JJA and CGSP14 all of which lack the *SPD0424-28* operon. Interestingly, this genomic island is absent in 50% pneumococcal strains including D39 but it has around 30% identity with the PTS system encoded by the *SPD0424-28* operon. The absence of a genomic island in these strains (D39, R6, TIGR4, Taiwan 19F, TCH8431 and 670-6B) might suggest that the *SPD0424-28* operon has a similar role in pathogenesis of these pneumococcal strains lacking the PTS-encoding genomic island studied by McAllister and co-workers (164).

Blast searches of SPD0424-28 revealed 30% sequence identity with pneumococcal proteins involved in cellobiose and lactose utilization. However, data of a transcriptome comparison of cellobiose vs glucose (Chapter 5) and lactose vs glucose (unpublished data) indicates no effect of these sugars on the expression of *SPD0424-28*, consistent with P*SPD0424-lacZ* expression data in the presence of different sugar sources (Fig. 6). This suggests that this operon is probably not involved in the metabolism of these sugars. Moreover, to determine the conservation of RokA in other streptococcal species, we blasted the sequence of RokA in different *Streptococci*. This revealed that RokA is absent in many *Streptococci* including *Streptococcus mutans*, *Streptococcus gordonii* and *Streptococcus pyogenes*. However, in other streptococci like *Streptococcus suis*, *Streptococcus gallolyticus* and *Streptococcus equi* a homolog can be found. In these *streptococci*, *rokA* is annotated as XylR, a xylose-dependent repressor. However, our β-galactosidase assays with P*SPD0424-lacZ* showed no effect of xylose on the expression of the *SPD0424-28* operon (Fig. 6). It might be that RokA interacts with structural analogs of different sugars in a natural host environment and one or several of which could be the substrate for RokA.

Acknowledgements

SS is supported by the HEC-NUFFIC overseas scholarship program by HEC (Higher Education commission) Pakistan and TGK was supported by the EU-funded Pneumopath project HEALTH-F3-2009-222983.

Chapter 7

Summary and general discussion

Streptococcus pneumoniae resides as a commensal in the nasopharynx (31). However, during favorable conditions, it spreads from the nasopharynx to different parts of the human body to cause pneumococcal infections (114). To survive and cause infection in these very different host environments, *S. pneumoniae* must be able to sense and adapt to the considerable variation in environmental conditions, including changes in the concentration of metal ions and availability of different carbon sources in host tissues. This thesis provides further knowledge on the effects of varying metal ions and carbohydrate availability on gene regulation and virulence of *S. pneumoniae*. Furthermore, this thesis describes the functional characterization of the transcriptional regulators AdcR, CopY, CelR, CelR-II and RokA of *S. pneumoniae* that respond to various environmental cues.

1. Role of trace metal ions in gene regulation and virulence of *S. pneumoniae*

Proper homeostasis of trace metals is vital for the survival of *S. pneumoniae*. For this purpose, *S. pneumoniae* possesses various trace metal ion-uptake and -efflux systems, which are tightly controlled by metal-dependent transcriptional regulators. Fig. 1 summarizes the metal-dependent transcriptional regulators and the genes that are controlled by these regulators in *S. pneumoniae*. These metal-dependent transcriptional regulators usually bind to a specific trace metal but some (like SczA, CopY and PsaR) have the ability to bind more than one kind of metal ion. Retaining proper concentrations of trace metals is important for bacteria to maintain proper metal ion homeostasis and virulence.

An important trace metal that *S. pneumoniae* faces in the environment is zinc. It is toxic to the cell at high concentrations and required for the proper functioning of various enzymes and the regulation of virulence and ribosomal genes (21, 128, 129, 185, 188, 210, 224). *S. pneumoniae* has dedicated separate systems for zinc-uptake (during zinc-limitation: the *adc* operon) and -efflux (during zinc-stress: the *czcD* gene), which are regulated independently via different transcriptional regulators, *i.e.* AdcR and SczA, respectively. Zinc-limitation affects the expression of the zinc uptake operon (*adc* operon) and several virulence genes, including those encoding pneumococcal histidine triad (Pht) proteins (PhtA, PhtB, PhtD and PhtE) and laminin binding protein (LmB/AdcRII) in *S. pneumoniae* (21, 70, 149, 185, 207, 224). In **Chapter 2**, we showed by using transcriptome analysis, *lacZ*-reporter studies, *in vitro* DNA binding experiments, and *in silico* operator predictions that AdcR directly represses the promoters of *adcRCBA*, *adcAII-phtD*, *phtA*, *phtB* and *phtE* in the presence of zinc. Moreover, AdcR can also function as an activator, since in the presence of

zinc it directly induces expression of *adh*, which encodes a zinc-containing alcohol dehydrogenase. On the other hand, a cation efflux system (CzcD) is responsible to cope with zinc toxicity and the expression of *czcD* is regulated by the TetR-family transcriptional regulator SczA (128).

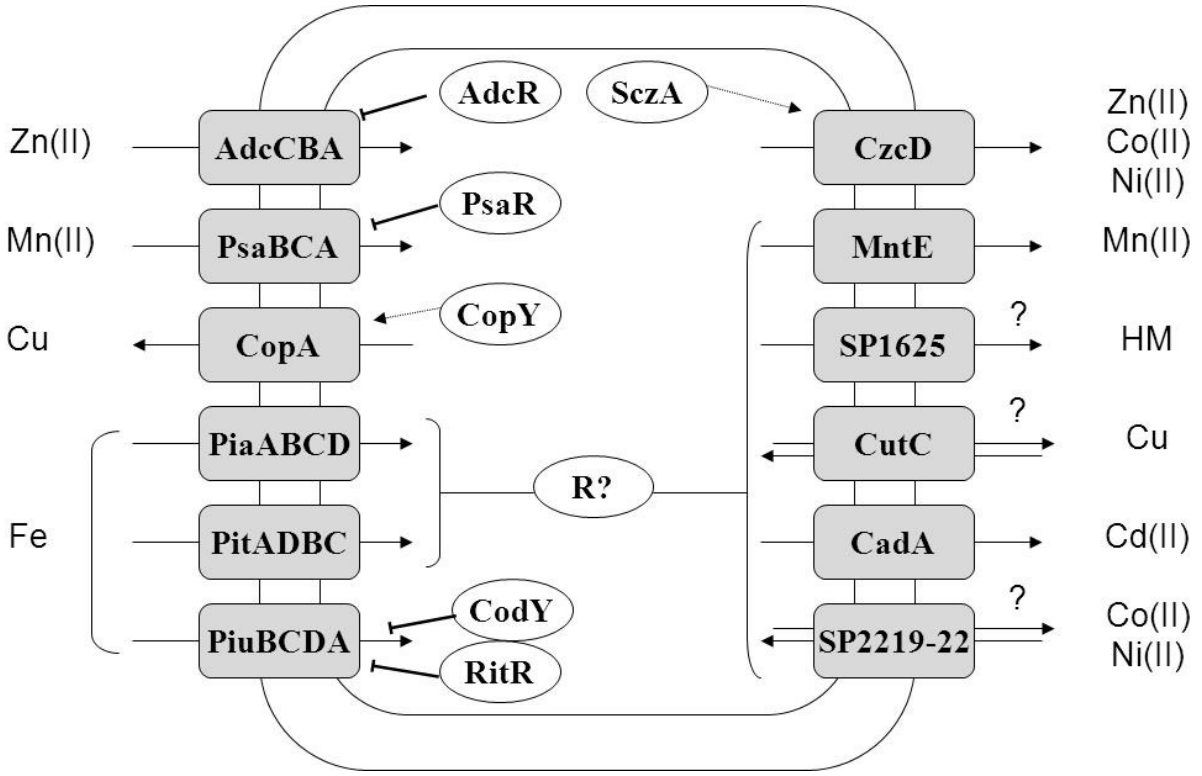


Figure 1: Summary of the metal transport genes that are encoded by the genome of D39.

Another important but toxic trace metal *S. pneumoniae* might encounter in its natural environment is copper. The genome of *S. pneumoniae* strain D39 possesses copper responsive genes, including the *cop* operon, and the *cutC*, *ctpC* and *ctpE* genes. The *cop* operon in *S. pneumoniae* is comprised of three genes (*copY*, *cupA* and *copA*) that are regulated by the copper-responsive transcriptional regulator CopY (**Chapter 3**). The expression of the *cop* operon is activated by copper and repressed by zinc. Moreover, susceptibility of *S. pneumoniae* to relatively high copper concentrations was increased due to deletion of *copA*, whereas little effect on the growth was observed due to a *cupA* mutation (225). The *cop* operon was also induced in the lungs and nasopharynx of intranasal infected mice and reduced virulence in pneumococcal strains lacking *copA* was observed in a mouse model of pneumonia (225). Interestingly, DNA microarray analysis with high and low copper ion concentrations only revealed upregulation of the *cop* operon and no effect on the expression of the *cutC*, *ctpC* and *ctpE* genes were observed. However, expression of these genes was

increased in lungs compared to blood (**Chapter 3**). Moreover, expression of *cutC* was also increased in the nasopharynx as compared to blood. This suggests the involvement of these genes in the virulence of *S. pneumoniae*, although the exact mechanism of their regulation remains unknown.

Since the virulence of *S. pneumoniae* displays a clear dependence on transition metal ion-uptake systems, there could be possibilities to use these systems as targets for anti-infectious therapies. This could be achieved by generating compounds that inhibit such systems, or vaccines based on these systems could be developed. One class of vaccine targets could be the Pht family of proteins, which is highly conserved in pneumococcal strains, and especially PhtD, which shows very little sequence variability (210). All Pht family proteins were able to elicit a high level protection in virulence settings (2, 22, 67, 89, 184, 267, 279). PhtD and E, as well as PcpA, even elicited an antibody response in humans (193), which might lead to decreased adhesion to epithelial cells as mediated by these proteins (95). PhtD and E were also found in an antigen screen using human and mice sera (22). PcpA also protects against sepsis and lung infection, but despite the fact that it has been reported to mediate adherence in the nasopharynx, not against colonization of the nose (78). Also LmB, which binds zinc and presumably has a similar role as AdcAII in *S. pneumoniae*, is conserved in all strains of *Streptococcus agalactiae* and might be a good vaccine candidate (203).

S. pneumoniae contains up to four zinc metalloproteases *i.e.* IgA, ZmpB, C and D. ZmpB is highly conserved among various serotypes and is involved in pneumococcal virulence. It is an interesting vaccine candidate, which offers protection against pneumococcal invasive disease in mice, especially when combined with other antigens. Zinc also reduces the toxicity of pneumolysin to cochlear hair cells *in vitro* due to an inhibition of toxin incorporation and aggregation into the plasma membrane, thus preventing calcium influx through the toxin pores, so this could also be a basis for therapy (74). Therefore, specific zinc concentrations might reduce the incidence of infectious diseases caused by *S. pneumonia*e.

The iron uptake proteins PiuA and PiaA are also good vaccine candidates. The recombinant PiuA and PiaA proteins increased the immunity in mice against systemic infections (38, 112). Both proteins have the ability to generate an antibody response in mice (112). Moreover, anti-PiaA and -PiuA polyclonal rabbit antibodies were able to bind to the surface of *S. pneumoniae*. Opsonophagocytosis of different pneumococcal serotypes by human neutrophil cell lines was increased by these antibodies (112). Although the role of iron has been already established in the virulence of *S. pneumoniae*, still the impact of iron on gene expression of *S. pneumoniae* needs more attention. Previously, the role of the orphan

response regulator (RitR) and the global nutritional transcriptional regulator (CodY) in the regulation of iron uptake system (PiuBCDA) was established (47, 251). However, the exact role of iron in the regulation of *piuBCDA* by RitR/CodY has not been explored yet. In addition to PiuBCDA, *S. pneumoniae* also encodes two more iron uptake systems (PiaABCD and PitADBC) (Fig. 1). The transcriptional regulators for these two iron uptake systems are also not known.

So far, inhibitory molecules have not been generated based on metal ion homeostasis systems, but such compounds could be well targeted at one of the zinc/copper systems. Thus, a proper understanding of trace metals in relation to the virulence of *S. pneumoniae* might open up new avenues for protection against infection.

2. Carbohydrate utilization in relation to virulence of *S. pneumoniae*

Another important environmental challenge *S. pneumoniae* might face is the presence or absence of various carbon sources in the host. Glucose is considered as a preferred energy source for *S. pneumoniae*, but the availability of glucose in other places inside the human body is low (228). This indicates the necessity for *S. pneumoniae* to utilize other sugar sources. **Chapters 4** and **5** highlight the effects of cellobiose on gene expression of *S. pneumoniae* and the functional characterization of the transcriptional regulators CelR and CelR-II that are responsive to cellobiose. Cellobiose is one of the β-glucosides that can be a potential source of energy for different streptococci including *S. pneumoniae* (29, 168, 275). However, it is unlikely that cellobiose can be present in the human host as a natural substrate. The extracellular matrix of mammalian tissues is rich in glycosaminoglycans that contain repeating units of β-linked disaccharides (119). *S. pneumoniae* can cleave host glucosaminoglycans by the activity of many glucosidases, thereby releasing useable carbon sources that are structural analogs of cellobiose (121). Probably, *S. pneumoniae* may encounter any of these structural analogs of cellobiose in the natural environment.

S. pneumoniae has at least two loci (the *cel* locus and the *cel-II* operon) that respond to cellobiose (**Chapter 4** & **5**). The *cel* locus, which transcribes into two transcriptional units from the promoter of *celA* and *celB*, is comprised of seven genes encoding a 6-phospho-β-glucosidase (CelA), a putative DNA-binding transcriptional regulator (CelR), the A, B and C domains of a cellobiose PTS EII^{Cel} (CelB, CelC, and CelD) and two hypothetical proteins. Previously, it has been shown that the *cel* locus is regulated by TCS08 (Two Component System 08) and is involved in cellobiose utilization in *S. pneumoniae* strain R6 (168).

However, in **Chapter 4**, deletion of *hk08* has no effect on the expression of P*celA* and P*celB* in D39 strain (223). This means that TCS08 probably regulates the *cel* locus in response to a different stimulus than the sugar source in strain D39. Moreover, despite the presence of a *cre* box in P*celA*, the *ccpA* mutation has no effect on the expression of the *cel* locus (223). Thus, most likely CelR is the only regulator involved in the regulation of the *cel* locus in response to the sugar source in strain D39. In addition, CelR mediated the expression of the *cel* locus by binding to a 22-bp regulatory site present in P*celA* and P*celB*. Interestingly, this binding site is highly conserved in different streptococci that have the *cel* locus. This suggests a similar fashion of regulation by CelR in these streptococci as well. However, 50% of the strains available at the KEGG website do not contain genes homologous to the *cel* locus, implying that the role of TCS08 lies outside regulation of the *cel* locus in these strains.

Transcriptome analysis in the presence of cellobiose revealed the presence of a second cellobiose-responsive operon (*cel-II* operon) (**Chapter 5**). The *cel-II* operon is comprised of four genes encoding a glycosyl hydrolase (*celA-II*) and a PTS system IICBA component (*celDCB-II*). Upstream of *celA-II*, a GntR-family transcriptional regulator, *celR-II* is located that acts as a transcriptional repressor of the *cel-II* operon in the absence of cellobiose. CelR-II mediates this repression by binding to a 20-bp operator site present in P*celA-II*. Moreover, deletion of *ccpA* also has no impact on expression of the *cel-II* operon. Thus, CelR-II directly controls the expression of the *cel-II* operon. The *cel-II* operon has high sequence homology with the *cel* locus and is highly conserved in all the pneumococcal strains available on the KEGG website. The deletion of *celR* decreased normal growth in the presence of cellobiose. However, there was no impact of *celDCB-II* deletion on normal growth in the presence of cellobiose. This shows the large importance of the *cel* locus in cellobiose utilization. Moreover, the exact role of the *cel-II* operon in *S. pneumoniae* D39 strain still remains elusive. It could play a role in the utilization of β-glucosides derived from host sugar polymers as discussed above.

Genes involved in cellobiose utilization are also important for the pathogenesis of *S. pneumoniae*, as in previous STM screens, CelR and CelD-II were found important for the virulence of *S. pneumoniae* (93, 137). Several pneumococcal strains, like AP200, ATCC700669-23F, G54-19F, Hungary-19A, JJA, P1031 and TCH8431/19A lack the *cel* locus, hence the contribution of the *cel* locus to virulence seems strain-specific. On the other hand, *celD-II* is highly conserved in all pneumococcal strains. Therefore, it will be interesting to explore the exact role of the *cel-II* locus in pathogenesis and physiology of *S. pneumoniae* in animal models.

It appears that cellobiose utilization also has a link with maltose utilization. On both time points of the transcriptome analysis with cellobiose, the expression of maltose utilization genes was altered. This altered expression might be due to upregulation of the *cel-II* operon. *celR-II* deletion also led to increased expression of maltose utilization genes, which clearly indicates a regulatory role of *celR-II* in maltose utilization. Maltose utilization in *S. pneumoniae* is known to be regulated by MalR (maltose-dependent repressor) (179, 180). Further research should reveal the exact relation of CelR-II and MalR in the regulation of maltose utilization genes.

The transcriptional regulator that is described in **Chapter 6** is RokA. RokA belongs to the ROK family of proteins that is comprised of transcriptional repressors, sugar kinases and uncharacterised open reading frames (ORFs) (54, 245). RokA acts as a transcriptional repressor of a putative carbohydrate utilization operon (*SPD0424-28*). This repression is mediated by direct binding of RokA to a 25-bp DNA AT-rich operator sequence that is present in P*SPD0424*. Moreover, transcriptome analysis with a *rokA* mutant revealed that the *SPD0424-28* operon is the only direct target of RokA. Despite the fact that the *SPD0424-28* operon has high homology with a lactose/cellobiose specific PTS, while RokA has high homology with XylR, we were not able to identify a possible co-factor that causes derepression of *SPD0424-28* operon by RokA. Further investigation is required to find the possible sugar that might act as an inducer of RokA.

The role of ROK family proteins in carbohydrate utilization has already been described in various bacteria including *Bacillus subtilis*, *Escherichia coli* and *S. pneumoniae* (64, 71, 120, 132, 194-196). Recently, a ROK transcriptional regulator that regulates operons encoding a PTS and sulfatase, all present on a genomic island, was found to be involved in the pathogenesis of *S. pneumoniae* strain WCH206 (164). This genomic island is not present in the D39 genome. However, the *SPD0424-28* operon has high homology with this PTS although a gene encoding a sulfatase is not present in the vicinity of *SPD0424-28*. The *SPD0424-28* operon could have the same role in strain D39 as the PTS encoded on the genomic island in strain WCH206. Thus a role of RokA and its target genes in the virulence of *S. pneumoniae* strain D39 might be well possible. As mentioned in **Chapter 6**, *S. pneumoniae* strain D39 encodes at least three ROK family kinases and it will be interesting to investigate the role of these kinases in carbohydrate utilization and virulence.

Concluding remarks

It is getting more and more clear that in order to understand details of pneumococcal infection one must take into account the environmental/nutritional conditions faced by *S. pneumoniae* inside the host. This thesis will help us to understand more about the effects of different environmental stimuli *i.e.* varying concentrations of metal ions and the availability of different carbohydrates on gene expression in the human pathogen *S. pneumoniae*. The genes responding to varying metal concentrations and to the presence of different carbohydrates are also linked with the virulence of *S. pneumoniae*. Therefore, understanding more about the regulation of these genes might help us to use them as a target to counteract the pathogenesis of *S. pneumoniae*.

References

References

1. **Abranches J, Chen YY, Burne RA.** 2004. Galactose metabolism by *Streptococcus mutans*. Appl. Environ. Microbiol. **70:**6047-6052.
2. **Adamou JE, Heinrichs JH, Erwin AL, Walsh W, Gayle T, Dormitzer M, Dagan R, Brewah YA, Barren P, Lathigra R, Langermann S, Koenig S, Johnson S.** 2001. Identification and characterization of a novel family of pneumococcal proteins that are protective against sepsis. Infect. Immun. **69:**949-958.
3. **Adamou JE, Wizemann TM, Barren P, Langermann S.** 1998. Adherence of *Streptococcus pneumoniae* to human bronchial epithelial cells (BEAS-2B). Infect. Immun. **66:**820-822.
4. **Albano M, Smits WK, Ho LT, Kraigher B, Mandic-Mulec I, Kuipers OP, Dubnau D.** 2005. The Rok protein of *Bacillus subtilis* represses genes for cell surface and extracellular functions. J. Bacteriol. **187:**2010-2019.
5. **Alloing G, Granadel C, Morrison DA, Claverys JP.** 1996. Competence pheromone, oligopeptide permease, and induction of competence in *Streptococcus pneumoniae*. Mol. Microbiol. **21:**471-478.
6. **Alteri CJ and Mobley HL.** 2012. *Escherichia coli* physiology and metabolism dictates adaptation to diverse host microenvironments. Curr. Opin. Microbiol. **15:**3-9.
7. **Amon J, Titgemeyer F, Burkovski A.** 2010. Common patterns - unique features: nitrogen metabolism and regulation in Gram-positive bacteria. FEMS Microbiol. Rev. **34:**588-605.
8. **Ando M, Manabe YC, Converse PJ, Miyazaki E, Harrison R, Murphy JR, Bishai WR.** 2003. Characterization of the role of the divalent metal ion-dependent transcriptional repressor MntR in the virulence of *Staphylococcus aureus*. Infect. Immun. **71:**2584-2590.
9. **Andreini C, Bertini I, Cavallaro G, Holliday GL, Thornton JM.** 2008. Metal ions in biological catalysis: from enzyme databases to general principles. J. Biol. Inorg. Chem. **13:**1205-1218.
10. **Aranda J, Garrido ME, Cortes P, Llagostera M, Barbe J.** 2008. Analysis of the protective capacity of three *Streptococcus suis* proteins induced under divalent-cation-limited conditions. Infect. Immun. **76:**1590-1598.
11. **Aranda J, Garrido ME, Fittipaldi N, Cortes P, Llagostera M, Gottschalk M, Barbe J.** 2010. The cation-uptake regulators AdcR and Fur are necessary for full virulence of *Streptococcus suis*. Vet. Microbiol. **144:**246-249.
12. **Aranda J, Garrido ME, Fittipaldi N, Cortes P, Llagostera M, Gottschalk M, Barbe J.** 2009. Protective capacities of cell surface-associated proteins of *Streptococcus suis* mutants deficient in divalent cation-uptake regulators. Microbiology **155:**1580-1587.
13. **Arredondo M and Nunez MT.** 2005. Iron and copper metabolism. Mol. Aspects Med. **26:**313-327.
14. **Austrian R.** 1999. The Pneumococcus at the Millennium: Not Down, Not Out. J. Infect. Dis. **179:**S338-S341.
15. **Avery OT, Macleod CM, McCarty M.** 1944. Studies on the chemical nature of the substance inducing transformation of pneumococcal types. Induction of transformation by a desoxyribonucleic acid fraction isolated from Pneumococcus type III. 1944. Mol. Med. **1:**344-365.
16. **Baerends RJS, Smits WK, de Jong A, Hamoen LW, Kok J, Kuipers OP.** 2004. Genome2D: a visualization tool for the rapid analysis of bacterial transcriptome data. Genome Biol. **5(5):**R37.
17. **Bailey TL and Elkan C.** 1994. Fitting a mixture model by expectation maximization to discover motifs in biopolymers. Proc. Int. Conf. Intell. Syst. Mol. Biol. **2:**28-36.
18. **Baker J, Sitthisak S, Sengupta M, Johnson M, Jayaswal RK, Morrissey JA.** 2010. Copper stress induces a global stress response in *Staphylococcus aureus* and represses *sae* and *agr* expression and biofilm formation. Appl. Environ. Microbiol. **76:**150-160.
19. **Balachandran P, Hollingshead SK, Paton JC, Briles DE.** 2001. The autolytic enzyme LytA of *Streptococcus pneumoniae* is not responsible for releasing pneumolysin. J. Bacteriol. **183:**3108-3116.
20. **Barabote RD and Saier MH,Jr.** 2005. Comparative genomic analyses of the bacterial phosphotransferase system. Microbiol. Mol. Biol. Rev. **69:**608-634.
21. **Bayle L, Chimalapati S, Schoehn G, Brown J, Vernet T, Durmort C.** 2011. Zinc uptake by *Streptococcus pneumoniae* depends on both AdcA and AdcAII and is essential for normal bacterial morphology and virulence. Mol. Microbiol. **82:**904-916.
22. **Beghetto E, Gargano N, Ricci S, Garufi G, Peppoloni S, Montagnani F, Oggioni M, Pozzi G, Felici F.** 2006. Discovery of novel *Streptococcus pneumoniae* antigens by screening a whole-genome lambda-display library. FEMS Microbiol. Lett. **262:**14-21.
23. **Belitsky BR, Kim HJ, Sonenshein AL.** 2004. CcpA-dependent regulation of *Bacillus subtilis* glutamate dehydrogenase gene expression. J. Bacteriol. **186:**3392-3398.
24. **Benton KA, Paton JC, Briles DE.** 1997. Differences in virulence for mice among *Streptococcus pneumoniae* strains of capsular types 2, 3, 4, 5, and 6 are not attributable to differences in pneumolysin production. Infect. Immun. **65:**1237-1244.
25. **Berry AM, Lock RA, Hansman D, Paton JC.** 1989. Contribution of autolysin to virulence of *Streptococcus pneumoniae*. Infect. Immun. **57:**2324-2330.

26. **Berry AM, Lock RA, Paton JC.** 1996. Cloning and characterization of nanB, a second *Streptococcus pneumoniae* neuraminidase gene, and purification of the NanB enzyme from recombinant *Escherichia coli*. J. Bacteriol. **178:**4854-4860.
27. **Berry AM and Paton JC.** 2000. Additive attenuation of virulence of *Streptococcus pneumoniae* by mutation of the genes encoding pneumolysin and other putative pneumococcal virulence proteins. Infect. Immun. **68:**133-140.
28. **Bethe G, Nau R, Wellmer A, Hakenbeck R, Reinert RR, Heinz HP, Zysk G.** 2001. The cell wall-associated serine protease PrtA: a highly conserved virulence factor of *Streptococcus pneumoniae*. FEMS Microbiol. Lett. **205:**99-104.
29. **Bidossi A, Mulas L, Decorosi F, Colomba L, Ricci S, Pozzi G, Deutscher J, Viti C, Oggioni MR.** 2012. A functional genomics approach to establish the complement of carbohydrate transporters in *Streptococcus pneumoniae*. PLoS One **7:**e33320.
30. **Blencowe DK and Morby AP.** 2003. Zn(II) metabolism in prokaryotes. FEMS Microbiol. Rev. **27:**291-311.
31. **Bogaert D, de GR, Hermans PW.** 2004. *Streptococcus pneumoniae* colonisation: the key to pneumococcal disease. Lancet Infect. Dis. **4:**144-154.
32. **Brandt EG, Hellgren M, Brinck T, Bergman T, Edholm O.** 2009. Molecular dynamics study of zinc binding to cysteines in a peptide mimic of the alcohol dehydrogenase structural zinc site. Phys. Chem. Chem. Phys. **11:**975-983.
33. **Bray BA, Sutcliffe IC, Harrington DJ.** 2009. Expression of the MtsA lipoprotein of *Streptococcus agalactiae* A909 is regulated by manganese and iron. Antonie Van Leeuwenhoek **95:**101-109.
34. **Brenot A, Weston BF, Caparon MG.** 2007. A PerR-regulated metal transporter (PmtA) is an interface between oxidative stress and metal homeostasis in *Streptococcus pyogenes*. Mol. Microbiol. **63:**1185-1196.
35. **Brown JS, Gilliland SM, Holden DW.** 2001. A *Streptococcus pneumoniae* pathogenicity island encoding an ABC transporter involved in iron uptake and virulence. Mol. Microbiol. **40:**572-585.
36. **Brown JS, Gilliland SM, Ruiz-Albert J, Holden DW.** 2002. Characterization of *pit*, a *Streptococcus pneumoniae* iron uptake ABC transporter. Infect. Immun. **70:**4389-4398.
37. **Brown JS and Holden DW.** 2002. Iron acquisition by Gram-positive bacterial pathogens. Microbes Infect. **4:**1149-1156.
38. **Brown JS, Ogunniyi AD, Woodrow MC, Holden DW, Paton JC.** 2001. Immunization with components of two iron uptake ABC transporters protects mice against systemic *Streptococcus pneumoniae* infection. Infect. Immun. **69:**6702-6706.
39. **Brown SA, Palmer KL, Whiteley M.** 2008. Revisiting the host as a growth medium. Nat. Rev. Microbiol. **6:**657-666.
40. **Buckwalter CM and King SJ.** 2012. Pneumococcal carbohydrate transport: food for thought Trends Microbiol. In press.
41. **Canvin JR, Marvin AP, Sivakumaran M, Paton JC, Boulnois GJ, Andrew PW, Mitchell TJ.** 1995. The role of pneumolysin and autolysin in the pathology of pneumonia and septicemia in mice infected with a type 2 pneumococcus. J. Infect. Dis. **172:**119-123.
42. **Cardozo DM, Nascimento-Carvalho CM, Souza FR, Silva NM.** 2006. Nasopharyngeal colonization and penicillin resistance among pneumococcal strains: a worldwide 2004 update Braz. J. Infect. Dis. **10:**293-304.
43. **Carvalho SM, Kloosterman TG, Kuipers OP, Neves AR.** 2011. CcpA ensures optimal metabolic fitness of *Streptococcus pneumoniae* D39. PLoS One **6:**e26707.
44. **Cassat JE and Skaar EP.** 2012. Metal ion acquisition in *Staphylococcus aureus*: overcoming nutritional immunity. Semin. Immunopathol. **34(2):**215-235.
45. **Castro R, Neves AR, Fonseca LL, Pool WA, Kok J, Kuipers OP, Santos H.** 2009. Characterization of the individual glucose uptake systems of *Lactococcus lactis*: mannose-PTS, cellobiose-PTS and the novel GlcU permease Mol. Microbiol. **71:**795-806.
46. **Catalani S, De PG, Mangili A, Apostoli P.** 2008. Metallic elements in lung tissues: results of a meta-analysis. Acta Biomed. **79 Suppl 1:**52-63.
47. **Caymaris S, Bootsma HJ, Martin B, Hermans PW, Prudhomme M, Claverys JP.** 2010. The global nutritional regulator CodY is an essential protein in the human pathogen *Streptococcus pneumoniae*. Mol. Microbiol. **78:**344-360.
48. **Chan PF, O'Dwyer KM, Palmer LM, Ambrad JD, Ingraham KA, So C, Lonetto MA, Biswas S, Rosenberg M, Holmes DJ, Zalacain M.** 2003. Characterization of a novel fucose-regulated promoter (PfcsK) suitable for gene essentiality and antibacterial mode-of-action studies in *Streptococcus pneumoniae*. J. Bacteriol. **185:**2051-2058.
49. **Chapuy-Regaud S, Ogunniyi AD, Diallo N, Huet Y, Desnottes JF, Paton JC, Escaich S, Trombe MC.** 2003. RegR, a global LacI/GalR family regulator, modulates virulence and competence in *Streptococcus pneumoniae*. Infect. Immun. **71:**2615-2625.

50. **Cobine P, Wickramasinghe WA, Harrison MD, Weber T, Solioz M, Dameron CT.** 1999. The *Enterococcus hirae* copper chaperone CopZ delivers copper(I) to the CopY repressor. FEBS Lett. **445:**27-30.
51. **Cobine PA, Jones CE, Dameron CT.** 2002. Role for zinc(II) in the copper(I) regulated protein CopY. J. Inorg. Biochem. **88:**192-196.
52. **Cochu A, Vadeboncoeur C, Moineau S, Frenette M.** 2003. Genetic and biochemical characterization of the phosphoenolpyruvate: glucose/mannose phosphotransferase system of *Streptococcus thermophiles*. Appl. Environ. Microbiol. **69:**5423-5432.
53. **Coleman JE.** 1998. Zinc enzymes. Curr. Opin. Chem. Biol. **2:**222-234.
54. **Conejo MS, Thompson SM, Miller BG.** 2010. Evolutionary bases of carbohydrate recognition and substrate discrimination in the ROK protein family. J. Mol. Evol. **70:**545-556.
55. **Cornelis P, Wei Q, Andrews SC, Vinckx T.** 2011. Iron homeostasis and management of oxidative stress response in bacteria. Metallomics **3:**540-549.
56. **Cote CK, Cvitkovitch D, Bleiweis AS, Honeyman AL.** 2000. A novel beta-glucoside-specific PTS locus from *Streptococcus mutans* that is not inhibited by glucose. Microbiology **146 (Pt 7):**1555-1563.
57. **Cron LE, Bootsma HJ, Noske N, Burghout P, Hammerschmidt S, Hermans PW.** 2009. Surface-associated lipoprotein PpmA of *Streptococcus pneumoniae* is involved in colonization in a strain-specific manner. Microbiology **155:**2401-2410.
58. **Cundell DR, Weiser JN, Shen J, Young A, Tuomanen EI.** 1995. Relationship between colonial morphology and adherence of *Streptococcus pneumoniae*. Infect. Immun. **63:**757-761.
59. **Dawid S, Roche AM, Weiser JN.** 2007. The *blp* bacteriocins of *Streptococcus pneumoniae* mediate intraspecies competition both in vitro and in vivo. Infect. Immun. **75:**443-451.
60. **de Prost N and Saumon G.** 2007. Glucose transport in the lung and its role in liquid movement Respir. Physiol. Neurobiol. **159:**331-337.
61. **de Ruyter PG, Kuipers OP, Beerthuyzen MM, van Alen-Boerrigter I, de Vos WM.** 1996. Functional analysis of promoters in the nisin gene cluster of *Lactococcus lactis*. J. Bacteriol. **178:**3434-3439.
62. **de Ruyter PG, Kuipers OP, de Vos WM.** 1996. Controlled gene expression systems for *Lactococcus lactis* with the food-grade inducer nisin. Appl. Environ. Microbiol. **62:**3662-3667.
63. **de Vos WM and Vaughan EE.** 1994. Genetics of lactose utilization in lactic acid bacteria. FEMS Microbiol. Rev. **15:**217-237.
64. **Decker K, Plumbridge J, Boos W.** 1998. Negative transcriptional regulation of a positive regulator: the expression of malT, encoding the transcriptional activator of the maltose regulon of *Escherichia coli*, is negatively controlled by Mlc. Mol. Microbiol. **27:**381-390.
65. **den Hengst CD, Curley P, Larsen R, Buist G, Nauta A, van Sinderen D, Kuipers OP, Kok J.** 2005. Probing direct interactions between CodY and the *oppD* promoter of *Lactococcus lactis*. J. Bacteriol. **187:**512-521.
66. **den Hengst CD, van Hijum SAFT, Geurts JM, Nauta A, Kok J, Kuipers OP.** 2005. The *Lactococcus lactis* CodY regulon: identification of a conserved cis-regulatory element. J. Biol. Chem. **280:**34332-34342.
67. **Denoel P, Godfroid F, Hermand P, Verlant V, Poolman J.** 2011. Combined protective effects of anti-PhtD and anti-pneumococcal polysaccharides. Vaccine **29(38):**6451-3.
68. **Deutscher J, Francke C, Postma PW.** 2006. How phosphotransferase system-related protein phosphorylation regulates carbohydrate metabolism in bacteria. Microbiol. Mol. Biol. Rev. **70:**939-1031.
69. **Dintilhac A, Alloing G, Granadel C, Claverys JP.** 1997. Competence and virulence of *Streptococcus pneumoniae*: Adc and PsaA mutants exhibit a requirement for Zn and Mn resulting from inactivation of putative ABC metal permeases. Mol. Microbiol. **25:**727-739.
70. **Dintilhac A and Claverys JP.** 1997. The *adc* locus, which affects competence for genetic transformation in *Streptococcus pneumoniae*, encodes an ABC transporter with a putative lipoprotein homologous to a family of streptococcal adhesins. Res. Microbiol. **148:**119-131.
71. **Dubeau MP, Poulin-Laprade D, Ghinet MG, Brzezinski R.** 2011. Properties of CsnR, the transcriptional repressor of the chitosanase gene, csnA, of *Streptomyces lividans*. J. Bacteriol. **193:**2441-2450.
72. **Elsner A, Kreikemeyer B, Braun-Kiewnick A, Spellerberg B, Buttaro BA, Podbielski A.** 2002. Involvement of Lsp, a member of the LraI-lipoprotein family in *Streptococcus pyogenes*, in eukaryotic cell adhesion and internalization. Infect. Immun. **70:**4859-4869.
73. **Finney LA and O'Halloran TV.** 2003. Transition metal speciation in the cell: insights from the chemistry of metal ion receptors. Science **300:**931-936.
74. **Franco-Vidal V, Beurg M, Darrouzet V, Bebear JP, Skinner LJ, Dulon D.** 2008. Zinc protection against pneumolysin toxicity on rat cochlear hair cells. Audiol. Neurootol. **13:**65-70.
75. **Gaballa A and Helmann JD.** 2003. *Bacillus subtilis* CPx-type ATPases: characterization of Cd, Zn, Co and Cu efflux systems. Biometals **16:**497-505.
76. **Garcia-Suarez MM, Vazquez F, Mendez FJ.** 2006. *Streptococcus pneumoniae* virulence factors and their clinical impact: An update. Enferm. Infecc. Microbiol. Clin. **24:**512-517.

77. **Giammarinaro P and Paton JC.** 2002. Role of RegM, a homologue of the catabolite repressor protein CcpA, in the virulence of *Streptococcus pneumoniae*. Infect. Immun. **70:**5454-5461.
78. **Glover DT, Hollingshead SK, Briles DE.** 2008. *Streptococcus pneumoniae* surface protein PcpA elicits protection against lung infection and fatal sepsis. Infect. Immun. **76:**2767-2776.
79. **Golynskiy MV, Gunderson WA, Hendrich MP, Cohen SM.** 2006. Metal binding studies and EPR spectroscopy of the manganese transport regulator MntR. Biochemistry **45:**15359-15372.
80. **Gorke B and Stulke J.** 2008. Carbon catabolite repression in bacteria: many ways to make the most out of nutrients. Nat. Rev. Microbiol. **6:**613-624.
81. **Gosink KK, Mann ER, Guglielmo C, Tuomanen EI, Masure HR.** 2000. Role of novel choline binding proteins in virulence of *Streptococcus pneumoniae*. Infect. Immun. **68:**5690-5695.
82. **Gough H, Luke GA, Beeley JA, Geddes DA.** 1996. Human salivary glucose analysis by high-performance ion-exchange chromatography and pulsed amperometric detection. Arch. Oral Biol. **41:**141-145.
83. **Graham AI, Hunt S, Stokes SL, Bramall N, Bunch J, Cox AG, McLeod CW, Poole RK.** 2009. Severe zinc depletion of *Escherichia coli*: roles for high affinity zinc binding by ZinT, zinc transport and zinc-independent proteins. J. Biol. Chem. **284:**18377-18389.
84. **Gu Y, Ding Y, Ren C, Sun Z, Rodionov DA, Zhang W, Yang S, Yang C, Jiang W.** 2010. Reconstruction of xylose utilization pathway and regulons in Firmicutes. BMC Genomics **11:**255.
85. **Guerra AJ and Giedroc DP.** 2012. Metal site occupancy and allosteric switching in bacterial metal sensor proteins. Arch. Biochem. Biophys. **519(2):**210-222.
86. **Gupta R, Shah P, Swiatlo E.** 2009. Differential gene expression in *Streptococcus pneumoniae* in response to various iron sources. Microb. Pathog. **47:**101-109.
87. **Gupta SD, Lee BT, Camakaris J, Wu HC.** 1995. Identification of *cutC* and *cutF* (*nlpE*) genes involved in copper tolerance in *Escherichia coli*. J. Bacteriol. **177:**4207-4215.
88. **Halfmann A, Hakenbeck R, Bruckner R.** 2007. A new integrative reporter plasmid for *Streptococcus pneumoniae*. FEMS Microbiol. Lett. **268:**217-224.
89. **Hamel J, Charland N, Pineau I, Ouellet C, Rioux S, Martin D, Brodeur BR.** 2004. Prevention of pneumococcal disease in mice immunized with conserved surface-accessible proteins. Infect. Immun. **72:**2659-2670.
90. **Hammer ND and Skaar EP.** 2011. Molecular mechanisms of *Staphylococcus aureus* iron acquisition. Annu. Rev. Microbiol. **65:**129-147.
91. **Hantke K.** 2005. Bacterial zinc uptake and regulators. Curr. Opin. Microbiol. **8:**196-202.
92. **Hantke K.** 2001. Bacterial zinc transporters and regulators. Biometals **14:**239-249.
93. **Hava DL and Camilli A.** 2002. Large-scale identification of serotype 4 *Streptococcus pneumoniae* virulence factors. Mol. Microbiol. **45:**1389-1406.
94. **Hendriksen WT, Bootsma HJ, Estevao S, Hoogenboezem T, de JA, de GR, Kuipers OP, Hermans PW.** 2008. CodY of *Streptococcus pneumoniae*: link between nutritional gene regulation and colonization. J. Bacteriol. **190:**590-601.
95. **Hendriksen WT, Bootsma HJ, van DA, Estevao S, Kuipers OP, de GR, Hermans PW.** 2009. Strain-specific impact of PsaR of *Streptococcus pneumoniae* on global gene expression and virulence. Microbiology **155:**1569-1579.
96. **Hendriksen WT, Kloosterman TG, Bootsma HJ, Estevao S, de GR, Kuipers OP, Hermans PW.** 2008. Site-specific contributions of glutamine-dependent regulator GlnR and GlnR-regulated genes to virulence of *Streptococcus pneumoniae*. Infect. Immun. **76:**1230-1238.
97. **Hermans PW, Adrian PV, Albert C, Estevao S, Hoogenboezem T, Luijendijk IH, Kamphausen T, Hammerschmidt S.** 2006. The streptococcal lipoprotein rotamase A (SlrA) is a functional peptidyl-prolyl isomerase involved in pneumococcal colonization. J. Biol. Chem. **281:**968-976.
98. **Higgins MA, Abbott DW, Boulanger MJ, Boraston AB.** 2009. Blood group antigen recognition by a solute-binding protein from a serotype 3 strain of *Streptococcus pneumoniae*. J. Mol. Biol. **388:**299-309.
99. **Honeyman AL and Curtiss R,3rd.** 2000. The mannitol-specific enzyme II (*mtlA*) gene and the *mtlR* gene of the PTS of *Streptococcus mutans*. Microbiology **146 (Pt 7):**1565-1572.
100. **Hoskins J, Alborn WE,Jr., Arnold J, Blaszczak LC, Burgett S, DeHoff BS, Estrem ST, Fritz L, Fu DJ, Fuller W, Geringer C, Gilmour R, Glass JS, Khoja H, Kraft AR, Lagace RE, LeBlanc DJ, Lee LN, Lefkowitz EJ, Lu J, Matsushima P, McAhren SM, McHenney M, McLeaster K, Mundy CW, Nicas TI, Norris FH, O'Gara M, Peery RB, Robertson GT, Rockey P, Sun PM, Winkler ME, Yang Y, Young-Bellido M, Zhao G, Zook CA, Baltz RH, Jaskunas SR, Rosteck PR,Jr., Skatrud PL, Glass JI.** 2001. Genome of the bacterium *Streptococcus pneumoniae* strain R6. J. Bacteriol. **183:**5709-5717.
101. **Hueck CJ and Hillen W.** 1995. Catabolite repression in *Bacillus subtilis*: a global regulatory mechanism for the gram-positive bacteria? Mol. Microbiol. **15:**395-401.
102. **Hvorup R, Chang AB, Saier MH,Jr.** 2003. Bioinformatic analyses of the bacterial L-ascorbate phosphotransferase system permease family. J. Mol. Microbiol. Biotechnol. **6:**191-205.

103. **Iyer R, Baliga NS, Camilli A.** 2005. Catabolite control protein A (CcpA) contributes to virulence and regulation of sugar metabolism in *Streptococcus pneumoniae*. J. Bacteriol. **187:**8340-8349.
104. **Iyer R and Camilli A.** 2007. Sucrose metabolism contributes to *in vivo* fitness of *Streptococcus pneumoniae*. Mol. Microbiol. **66:**1-13.
105. **Jacobsen FE, Kazmierczak KM, Lisher JP, Winkler ME, Giedroc DP.** 2011. Interplay between manganese and zinc homeostasis in the human pathogen *Streptococcus pneumoniae*. Metallomics. **3:**38-41.
106. **Jakubovics NS and Jenkinson HF.** 2001. Out of the iron age: new insights into the critical role of manganese homeostasis in bacteria. Microbiology **147:**1709-1718.
107. **Jakubovics NS and Valentine RA.** 2009. A new direction for manganese homeostasis in bacteria: identification of a novel efflux system in *Streptococcus pneumoniae*. Mol. Microbiol. **72:**1-4.
108. **Jedrzejas MJ.** 2004. Extracellular virulence factors of *Streptococcus pneumoniae*. Front. Biosci. **9:**891-914.
109. **Jeong JK, Kwon O, Lee YM, Oh DB, Lee JM, Kim S, Kim EH, Le TN, Rhee DK, Kang HA.** 2009. Characterization of the *Streptococcus pneumoniae* BgaC protein as a novel surface beta-galactosidase with specific hydrolysis activity for the Galbeta1-3GlcNAc moiety of oligosaccharides. J. Bacteriol. **191:**3011-3023.
110. **Johnston JW, Briles DE, Myers LE, Hollingshead SK.** 2006. Mn^{2+}-dependent regulation of multiple genes in *Streptococcus pneumoniae* through PsaR and the resultant impact on virulence. Infect. Immun. **74:**1171-1180.
111. **Johnston JW, Myers LE, Ochs MM, Benjamin WH,Jr., Briles DE, Hollingshead SK.** 2004. Lipoprotein PsaA in virulence of *Streptococcus pneumoniae*: surface accessibility and role in protection from superoxide. Infect. Immun. **72:**5858-5867.
112. **Jomaa M, Terry S, Hale C, Jones C, Dougan G, Brown J.** 2006. Immunization with the iron uptake ABC transporter proteins PiaA and PiuA prevents respiratory infection with *Streptococcus pneumoniae*. Vaccine **24:**5133-5139.
113. **Kadioglu A, Gingles NA, Grattan K, Kerr A, Mitchell TJ, Andrew PW.** 2000. Host cellular immune response to pneumococcal lung infection in mice. Infect. Immun. **68:**492-501.
114. **Kadioglu A, Weiser JN, Paton JC, Andrew PW.** 2008. The role of *Streptococcus pneumoniae* virulence factors in host respiratory colonization and disease. Nat. Rev. Microbiol. **6:**288-301.
115. **Kaufman GE and Yother J.** 2007. CcpA-dependent and -independent control of beta-galactosidase expression in *Streptococcus pneumoniae* occurs via regulation of an upstream phosphotransferase system-encoding operon. J. Bacteriol. **189:**5183-5192.
116. **Kazmierczak KM, Wayne KJ, Rechtsteiner A, Winkler ME.** 2009. Roles of rel(Spn) in stringent response, global regulation and virulence of serotype 2 *Streptococcus pneumoniae* D39. Mol. Microbiol. **72:**590-611.
117. **Kehl-Fie TE and Skaar EP.** 2010. Nutritional immunity beyond iron: a role for manganese and zinc. Curr. Opin. Chem. Biol. **14:**218-224.
118. **Kershaw CJ, Brown NL, Constantinidou C, Patel MD, Hobman JL.** 2005. The expression profile of *Escherichia coli* K-12 in response to minimal, optimal and excess copper concentrations. Microbiology **151:**1187-1198.
119. **Kilic AO, Tao L, Zhang Y, Lei Y, Khammanivong A, Herzberg MC.** 2004. Involvement of *Streptococcus gordonii* beta-glucoside metabolism systems in adhesion, biofilm formation, and *in vivo* gene expression. J. Bacteriol. **186:**4246-4253.
120. **Kimata K, Inada T, Tagami H, Aiba H.** 1998. A global repressor (Mlc) is involved in glucose induction of the *ptsG* gene encoding major glucose transporter in *Escherichia coli*. Mol. Microbiol. **29:**1509-1519.
121. **King SJ.** 2010. Pneumococcal modification of host sugars: a major contributor to colonization of the human airway? Mol. Oral Microbiol. **25:**15-24.
122. **King SJ, Hippe KR, Weiser JN.** 2006. Deglycosylation of human glycoconjugates by the sequential activities of exoglycosidases expressed by *Streptococcus pneumoniae*. Mol. Microbiol. **59:**961-974.
123. **Klein JS and Lewinson O.** 2011. Bacterial ATP-driven transporters of transition metals: physiological roles, mechanisms of action, and roles in bacterial virulence. Metallomics **3:**1098-1108.
124. **Kliegman JI, Griner SL, Helmann JD, Brennan RG, Glasfeld A.** 2006. Structural basis for the metal-selective activation of the manganese transport regulator of *Bacillus subtilis*. Biochemistry **45:**3493-3505.
125. **Kloosterman TG, Bijlsma JJE, Kok J, Kuipers OP.** 2006. To have neighbour's fare: extending the molecular toolbox for *Streptococcus pneumoniae*. Microbiology **152:**351-359.
126. **Kloosterman TG, Hendriksen WT, Bijlsma JJ, Bootsma HJ, van Hijum SA, Kok J, Hermans PW, Kuipers OP.** 2006. Regulation of glutamine and glutamate metabolism by GlnR and GlnA in *Streptococcus pneumoniae*. J. Biol. Chem. **281:**25097-25109.
127. **Kloosterman TG and Kuipers OP.** 2011. Regulation of arginine acquisition and virulence gene expression in the human pathogen *Streptococcus pneumoniae* by transcription regulators ArgR1 and AhrC. J. Biol. Chem. **286:**44594-44605.

128. **Kloosterman TG, van der Kooi-Pol MM, Bijlsma JJ, Kuipers OP.** 2007. The novel transcriptional regulator SczA mediates protection against $Zn^{(2+)}$ stress by activation of the Zn-resistance gene *czcD* in *Streptococcus pneumoniae*. Mol. Microbiol. **65:**1049-1063.
129. **Kloosterman TG, Witwicki RM, van der Kooi-Pol MM, Bijlsma JJ, Kuipers OP.** 2008. Opposite effects of Mn^{2+} and Zn^{2+} on PsaR-mediated expression of the virulence genes *pcpA*, *prtA*, and *psaBCA* of *Streptococcus pneumoniae*. J. Bacteriol. **190:**5382-5393.
130. **Klugman KP, Chien YW, Madhi SA.** 2009. Pneumococcal pneumonia and influenza: a deadly combination. Vaccine **27 Suppl 3:**C9-C14.
131. **Kornberg HL.** 2001. Routes for fructose utilization by *Escherichia coli*. J. Mol. Microbiol. Biotechnol. **3:**355-359.
132. **Kreuzer P, Gartner D, Allmansberger R, Hillen W.** 1989. Identification and sequence analysis of the *Bacillus subtilis* W23 *xylR* gene and *xyl* operator. J. Bacteriol. **171:**3840-3845.
133. **Kuipers OP, Beerthuyzen MM, de Ruyter PG, Luesink EJ, de Vos WM.** 1995. Autoregulation of nisin biosynthesis in *Lactococcus lactis* by signal transduction. J. Biol. Chem. **270:**27299-27304.
134. **Kuipers OP, Ruyter PG, Kleerebezem M, Vos WM.** 1998. Quorum sensing controlled gene expression in lactic acid bacteria. J. Biotechnol. **64:**15-21.
135. **Lampe DJ, Churchill ME, Robertson HM.** 1996. A purified mariner transposase is sufficient to mediate transposition *in vitro*. EMBO J. **15:**5470-5479.
136. **Lanie JA, Ng WL, Kazmierczak KM, Andrzejewski TM, Davidsen TM, Wayne KJ, Tettelin H, Glass JI, Winkler ME.** 2007. Genome sequence of Avery's virulent serotype 2 strain D39 of *Streptococcus pneumoniae* and comparison with that of unencapsulated laboratory strain R6. J. Bacteriol. **189:**38-51.
137. **Lau GW, Haataja S, Lonetto M, Kensit SE, Marra A, Bryant AP, McDevitt D, Morrison DA, Holden DW.** 2001. A functional genomic analysis of type 3 *Streptococcus pneumoniae* virulence. Mol. Microbiol. **40:**555-571.
138. **Le Breton Y, Pichereau V, Sauvageot N, Auffray Y, Rince A.** 2005. Maltose utilization in *Enterococcus faecalis*. J. Appl. Microbiol. **98:**806-813.
139. **Lee MR, Bae SM, Kim TS, Lee KJ.** 2006. Proteomic analysis of protein expression in *Streptococcus pneumoniae* in response to temperature shift. J. Microbiol. **44:**375-382.
140. **Leenhouts K, Buist G, Bolhuis A, ten Berge A, Kiel J, Mierau I, Dabrowska M, Venema G, Kok J.** 1996. A general system for generating unlabelled gene replacements in bacterial chromosomes. Mol. Gen. Genet. **253:**217-224.
141. **Leenhouts K, Venema G, Kok J.** 1998. A lactococcal pWV01 based integration toolbox for bacteria. Methods in Cell Science **20:**35-50.
142. **Lengeler JW and Jahreis K.** 2009. Bacterial PEP-dependent carbohydrate: phosphotransferase systems couple sensing and global control mechanisms. Contrib. Microbiol. **16:**65-87.
143. **Lewis K.** 2000. Programmed death in bacteria Microbiol. Mol. Biol. Rev. **64:**503-514.
144. **Li S, Kelly SJ, Lamani E, Ferraroni M, Jedrzejas MJ.** 2000. Structural basis of hyaluronan degradation by *Streptococcus pneumoniae* hyaluronate lyase. EMBO J. **19:**1228-1240.
145. **Linke C, Caradoc-Davies TT, Young PG, Proft T, Baker EN.** 2009. The laminin-binding protein Lbp from *Streptococcus pyogenes* is a zinc receptor. J. Bacteriol. **191:**5814-5823.
146. **Liu T, Ramesh A, Ma Z, Ward SK, Zhang L, George GN, Talaat AM, Sacchettini JC, Giedroc DP.** 2007. CsoR is a novel *Mycobacterium tuberculosis* copper-sensing transcriptional regulator. Nat. Chem. Biol. **3:**60-68.
147. **Llull D, Lopez R, Garcia E.** 2001. Genetic bases and medical relevance of capsular polysaccharide biosynthesis in pathogenic streptococci. Curr. Mol. Med. **1:**475-491.
148. **Loisel E, Chimalapati S, Bougault C, Imberty A, Gallet B, Di Guilmi AM, Brown J, Vernet T, Durmort C.** 2011. Biochemical characterization of the histidine triad protein PhtD as a cell surface zinc-binding protein of pneumococcus. Biochemistry **50(17):**3551-3558.
149. **Loisel E, Jacquamet L, Serre L, Bauvois C, Ferrer JL, Vernet T, Di Guilmi AM, Durmort C.** 2008. AdcAII, A New Pneumococcal Zn-Binding Protein Homologous with ABC Transporters: Biochemical and Structural Analysis. J. Mol. Biol. **381(3):**594-606.
150. **Long AD, Mangalam HJ, Chan BY, Tolleri L, Hatfield GW, Baldi P.** 2001. Improved statistical inference from DNA microarray data using analysis of variance and a Bayesian statistical framework. Analysis of global gene expression in *Escherichia coli* K12. J. Biol. Chem. **276:**19937-19944.
151. **Loo CY, Mitrakul K, Voss IB, Hughes CV, Ganeshkumar N.** 2003. Involvement of the *adc* operon and manganese homeostasis in *Streptococcus gordonii* biofilm formation. J. Bacteriol. **185:**2887-2900.
152. **Lopez R, Gonzalez MP, Garcia E, Garcia JL, Garcia P.** 2000. Biological roles of two new murein hydrolases of *Streptococcus pneumoniae* representing examples of module shuffling. Res. Microbiol. **151:**437-443.

153. **Lortie LA, Pelletier M, Vadeboncoeur C, Frenette M.** 2000. The gene encoding IIAB(Man)L in *Streptococcus salivarius* is part of a tetracistronic operon encoding a phosphoenolpyruvate: mannose/glucose phosphotransferase system. Microbiology **146 (Pt 3):**677-685.
154. **Lulko AT, Buist G, Kok J, Kuipers OP.** 2007. Transcriptome analysis of temporal regulation of carbon metabolism by CcpA in *Bacillus subtilis* reveals additional target genes. J. Mol. Microbiol. Biotechnol. **12:**82-95.
155. **Luo R, Mann B, Lewis WS, Rowe A, Heath R, Stewart ML, Hamburger AE, Sivakolundu S, Lacy ER, Bjorkman PJ, Tuomanen E, Kriwacki RW.** 2005. Solution structure of choline binding protein A, the major adhesin of *Streptococcus pneumoniae*. EMBO J. **24:**34-43.
156. **Macomber L and Imlay JA.** 2009. The iron-sulfur clusters of dehydratases are primary intracellular targets of copper toxicity. Proc. Natl. Acad. Sci. U. S. A. **106:**8344-8349.
157. **Magnani D, Barre O, Gerber SD, Solioz M.** 2008. Characterization of the CopR regulon of *Lactococcus lactis* IL1403. J. Bacteriol. **190:**536-545.
158. **Margolis E, Yates A, Levin BR.** 2010. The ecology of nasal colonization of *Streptococcus pneumoniae*, *Haemophilus influenzae* and *Staphylococcus aureus*: the role of competition and interactions with host's immune response. BMC Microbiol. **10:**59.
159. **Marion C, Aten AE, Woodiga SA, King SJ.** 2011. Identification of an ATPase, MsmK, which energizes multiple carbohydrate ABC transporters in *Streptococcus pneumoniae*. Infect. Immun. **79:**4193-4200.
160. **Marion C, Burnaugh AM, Woodiga SA, King SJ.** 2011. Sialic acid transport contributes to pneumococcal colonization. Infect. Immun. **79:**1262-1269.
161. **Marra A, Lawson S, Asundi JS, Brigham D, Hromockyj AE.** 2002. *In vivo* characterization of the *psa* genes from *Streptococcus pneumoniae* in multiple models of infection. Microbiology **148:**1483-1491.
162. **Martin B, Prudhomme M, Alloing G, Granadel C, Claverys JP.** 2000. Cross-regulation of competence pheromone production and export in the early control of transformation in *Streptococcus pneumoniae*. Mol. Microbiol. **38:**867-878.
163. **Maruyama Y, Nakamichi Y, Itoh T, Mikami B, Hashimoto W, Murata K.** 2009. Substrate specificity of streptococcal unsaturated glucuronyl hydrolases for sulfated glycosaminoglycan J. Biol. Chem. **284:**18059-18069.
164. **McAllister LJ, Ogunniyi AD, Stroeher UH, Paton JC.** 2012. Contribution of a genomic accessory region encoding a putative cellobiose phosphotransferase system to virulence of *Streptococcus pneumoniae*. PLoS One **7:**e32385.
165. **McAllister LJ, Tseng HJ, Ogunniyi AD, Jennings MP, McEwan AG, Paton JC.** 2004. Molecular analysis of the psa permease complex of *Streptococcus pneumoniae*. Mol. Microbiol. **53:**889-901.
166. **McDevitt CA, Ogunniyi AD, Valkov E, Lawrence MC, Kobe B, McEwan AG, Paton JC.** 2011. A molecular mechanism for bacterial susceptibility to zinc. PLoS. Pathog. **7(11):**e1002357.
167. **McGeer A and Low DE.** 2003. Is resistance futile? Nat. Med. **9:**390-392.
168. **McKessar SJ and Hakenbeck R.** 2007. The two-component regulatory system TCS08 is involved in cellobiose metabolism of *Streptococcus pneumoniae* R6. J. Bacteriol. **189:**1342-1350.
169. **Melin M, Di PE, Tikkanen L, Jarva H, Neyt C, Kayhty H, Meri S, Poolman J, Vakevainen M.** 2010. Interaction of pneumococcal histidine triad proteins with human complement. Infect. Immun. **78:**2089-2098.
170. **Miethke M and Marahiel MA.** 2007. Siderophore-based iron acquisition and pathogen control. Microbiol. Mol. Biol. Rev. **71:**413-451.
171. **Mitchell AM and Mitchell TJ.** 2010. *Streptococcus pneumoniae*: virulence factors and variation. Clin. Microbiol. Infect. **16:**411-418.
172. **Mitchell TJ.** 2003. The pathogenesis of streptococcal infections: from tooth decay to meningitis. Nat. Rev. Microbiol. **1:**219-230.
173. **Mitrakul K, Loo CY, Gyurko C, Hughes CV, Ganeshkumar N.** 2005. Mutational analysis of the *adcCBA* genes in *Streptococcus gordonii* biofilm formation. Oral Microbiol. Immunol. **20:**122-127.
174. **Moore CM and Helmann JD.** 2005. Metal ion homeostasis in *Bacillus subtilis*. Curr. Opin. Microbiol. **8:**188-195.
175. **Nairz M, Schroll A, Sonnweber T, Weiss G.** 2010. The struggle for iron - a metal at the host-pathogen interface Cell. Microbiol. **12:**1691-1702.
176. **Neef J, Farshchi A,V, Kim KS, Kuipers OP, Bijlsma JJ.** 2011. Deletion of a cation transporter promotes lysis in *Streptococcus pneumoniae*. Infect. Immun. **79(6):**2314-2323.
177. **Neelakanta G, Sankar TS, Schnetz K.** 2009. Characterization of a beta-glucoside operon (*bgc*) prevalent in septicemic and uropathogenic *Escherichia coli* strains. Appl. Environ. Microbiol. **75:**2284-2293.
178. **Nies DH.** 2003. Efflux-mediated heavy metal resistance in prokaryotes. FEMS Microbiol. Rev. **27:**313-339.
179. **Nieto C, Espinosa M, Puyet A.** 1997. The maltose/maltodextrin regulon of *Streptococcus pneumoniae*. Differential promoter regulation by the transcriptional repressor MalR. J. Biol. Chem. **272:**30860-30865.

180. **Nieto C, Puyet A, Espinosa M.** 2001. MalR-mediated regulation of the *Streptococcus pneumoniae malMP* operon at promoter PM. Influence of a proximal divergent promoter region and competition between MalR and RNA polymerase proteins. J. Biol. Chem. **276:**14946-14954.
181. **Novichkov PS, Laikova ON, Novichkova ES, Gelfand MS, Arkin AP, Dubchak I, Rodionov DA.** 2010. RegPrecise: a database of curated genomic inferences of transcriptional regulatory interactions in prokaryotes. Nucleic Acids Res. **38:**D111-8.
182. **Obaro S and Adegbola R.** 2002. The pneumococcus: carriage, disease and conjugate vaccines. J. Med. Microbiol. **51:**98-104.
183. **Ogunniyi AD, Folland RL, Briles DE, Hollingshead SK, Paton JC.** 2000. Immunization of mice with combinations of pneumococcal virulence proteins elicits enhanced protection against challenge with *Streptococcus pneumoniae*. Infect. Immun. **68:**3028-3033.
184. **Ogunniyi AD, Grabowicz M, Briles DE, Cook J, Paton JC.** 2007. Development of a vaccine against invasive pneumococcal disease based on combinations of virulence proteins of *Streptococcus pneumoniae*. Infect. Immun. **75:**350-357.
185. **Ogunniyi AD, Grabowicz M, Mahdi LK, Cook J, Gordon DL, Sadlon TA, Paton JC.** 2009. Pneumococcal histidine triad proteins are regulated by the Zn^{2+}-dependent repressor AdcR and inhibit complement deposition through the recruitment of complement factor. H. FASEB J. **23:**731-738.
186. **Old LA, Lowes S, Russell RR.** 2006. Genomic variation in *Streptococcus mutans*: deletions affecting the multiple pathways of beta-glucoside metabolism. Oral Microbiol. Immunol. **21:**21-27.
187. **Outten CE and O'Halloran TV.** 2001. Femtomolar sensitivity of metalloregulatory proteins controlling zinc homeostasis. Science **292:**2488-2492.
188. **Panina EM, Mironov AA, Gelfand MS.** 2003. Comparative genomics of bacterial zinc regulons: enhanced ion transport, pathogenesis, and rearrangement of ribosomal proteins. Proc. Natl. Acad. Sci. U. S. A. **100:**9912-9917.
189. **Papp-Wallace KM and Maguire ME.** 2006. Manganese Transport and the Role of Manganese in Virulence. Annu. Rev. Microbiol. **60:**187-209.
190. **Pericone CD, Park S, Imlay JA, Weiser JN.** 2003. Factors contributing to hydrogen peroxide resistance in *Streptococcus pneumoniae* include pyruvate oxidase (SpxB) and avoidance of the toxic effects of the fenton reaction. J. Bacteriol. **185:**6815-6825.
191. **Pettigrew MM, Gent JF, Revai K, Patel JA, Chonmaitree T.** 2008. Microbial interactions during upper respiratory tract infections. Emerg. Infect. Dis. **14:**1584-1591.
192. **Philips BJ, Meguer JX, Redman J, Baker EH.** 2003. Factors determining the appearance of glucose in upper and lower respiratory tract secretions. Intensive Care Med. **29:**2204-2210.
193. **Pichichero ME, Kaur R, Casey JR, Xu Q, Almudevar A, Ochs M.** 2012. Antibody response to *Streptococcus pneumoniae* proteins PhtD, LytB, PcpA, PhtE and Ply after nasopharyngeal colonization and acute otitis media in children. Hum. Vaccin Immunother. **8(6):**799-805.
194. **Plumbridge J.** 1995. Co-ordinated regulation of amino sugar biosynthesis and degradation: the NagC repressor acts as both an activator and a repressor for the transcription of the glmUS operon and requires two separated NagC binding sites. EMBO J. **14:**3958-3965.
195. **Plumbridge J and Pellegrini O.** 2004. Expression of the chitobiose operon of *Escherichia coli* is regulated by three transcription factors: NagC, ChbR and CAP Mol. Microbiol. **52:**437-449.
196. **Plumbridge JA.** 1991. Repression and induction of the nag regulon of *Escherichia coli* K-12: the roles of *nagC* and *nagA* in maintenance of the uninduced state. Mol. Microbiol. **5:**2053-2062.
197. **Portmann R, Poulsen KR, Wimmer R, Solioz M.** 2006. CopY-like copper inducible repressors are putative 'winged helix' proteins. Biometals **19:**61-70.
198. **Posfay-Barbe KM, Galetto-Lacour A, Grillet S, Ochs MM, Brookes RH, Kraehenbuhl JD, Cevey-Macherel M, Gehri M, Gervaix A, Siegrist CA.** 2011. Immunity to pneumococcal surface proteins in children with community-acquired pneumonia: a distinct pattern of responses to pneumococcal choline-binding protein A. Clin. Microbiol. Infect. **17(8):**1232-1238.
199. **Postma PW, Lengeler JW, Jacobson GR.** 1993. Phosphoenolpyruvate:carbohydrate phosphotransferase systems of bacteria. Microbiol. Rev. **57:**543-594.
200. **Potter AJ, Kidd SP, McEwan AG, Paton JC.** 2010. The MerR/NmlR family transcription factor of *Streptococcus pneumoniae* responds to carbonyl stress and modulates hydrogen peroxide production. J. Bacteriol. **192:**4063-4066.
201. **Preston JA and Dockrell DH.** 2008. Virulence factors in pneumococcal respiratory pathogenesis. Future Microbiol. **3:**205-221.
202. **Puyet A, Ibanez AM, Espinosa M.** 1993. Characterization of the *Streptococcus pneumoniae* maltosaccharide regulator MalR, a member of the LacI-GalR family of repressors displaying distinctive genetic features. J. Biol. Chem. **268:**25402-25408.

203. **Ragunathan P, Spellerberg B, Ponnuraj K.** 2009. Structure of laminin-binding adhesin (Lmb) from *Streptococcus agalactiae*. Acta Crystallogr. D Biol. Crystallogr. **65:**1262-1269.
204. **Regev-Yochay G, Lipsitch M, Basset A, Rubinstein E, Dagan R, Raz M, Malley R.** 2009. The pneumococcal pilus predicts the absence of *Staphylococcus aureus* co-colonization in pneumococcal carriers. Clin. Infect. Dis. **48:**760-763.
205. **Regev-Yochay G, Trzcinski K, Thompson CM, Malley R, Lipsitch M.** 2006. Interference between *Streptococcus pneumoniae* and *Staphylococcus aureus*: *In vitro* hydrogen peroxide-mediated killing by *Streptococcus pneumoniae*. J. Bacteriol. **188:**4996-5001.
206. **Reyes A, Leiva A, Cambiazo V, Mendez MA, Gonzalez M.** 2006. Cop-like operon: structure and organization in species of the Lactobacillale order. Biol. Res. **39:**87-93.
207. **Reyes-Caballero H, Guerra AJ, Jacobsen FE, Kazmierczak KM, Cowart D, Koppolu UM, Scott RA, Winkler ME, Giedroc DP.** 2010. The metalloregulatory zinc site in *Streptococcus pneumoniae* AdcR, a zinc-activated MarR family repressor. J. Mol. Biol. **403:**197-216.
208. **Reyes-Jara A, Latorre M, Lopez G, Bourgogne A, Murray BE, Cambiazo V, Gonzalez M.** 2010. Genome-wide transcriptome analysis of the adaptive response of *Enterococcus faecalis* to copper exposure. Biometals **23:**1105-1112.
209. **Riboldi-Tunnicliffe A, Isaacs NW, Mitchell TJ.** 2005. 1.2 Angstroms crystal structure of the *S. pneumoniae* PhtA histidine triad domain a novel zinc binding fold. FEBS Lett. **579:**5353-5360.
210. **Rioux S, Neyt C, Di PE, Turpin L, Charland N, Labbe S, Mortier MC, Mitchell TJ, Feron C, Martin D, Poolman JT.** 2010. Transcriptional regulation, occurrence and putative role of the Pht Family of *Streptococcus pneumoniae*. Microbiology **157(Pt 2):**336-348.
211. **Robillard GT and Broos J.** 1999. Structure/function studies on the bacterial carbohydrate transporters, enzymes II, of the phosphoenolpyruvate-dependent phosphotransferase system. Biochim. Biophys. Acta **1422:**73-104.
212. **Rosch JW, Gao G, Ridout G, Wang YD, Tuomanen EI.** 2009. Role of the manganese efflux system mntE for signalling and pathogenesis in *Streptococcus pneumoniae*. Mol. Microbiol. **72:**12-25.
213. **Rosch JW, Sublett J, Gao G, Wang YD, Tuomanen EI.** 2008. Calcium efflux is essential for bacterial survival in the eukaryotic host. Mol. Microbiol. **70:**435-444.
214. **Rosenow C, Maniar M, Trias J.** 1999. Regulation of the alpha-galactosidase activity in *Streptococcus pneumoniae*: characterization of the raffinose utilization system. Genome Res. **9:**1189-1197.
215. **Rubins JB, Paddock AH, Charboneau D, Berry AM, Paton JC, Janoff EN.** 1998. Pneumolysin in pneumococcal adherence and colonization. Microb. Pathog. **25:**337-342.
216. **Ryan KJ and Ray CG.** 2004. Sherris Medical Microbiology. McGraw Hill, USA.
217. **Saier MH„Jr. and Reizer J.** 1992. Proposed uniform nomenclature for the proteins and protein domains of the bacterial phosphoenolpyruvate: sugar phosphotransferase system. J. Bacteriol. **174:**1433-1438.
218. **Schaible UE and Kaufmann SH.** 2004. Iron and microbial infection. Nat. Rev. Microbiol. **2:**946-953.
219. **Schmittgen TD and Livak KJ.** 2008. Analyzing real-time PCR data by the comparative C(T) method. Nat. Protoc. **3:**1101-1108.
220. **Schwan WR, Warrener P, Keunz E, Stover CK, Folger KR.** 2005. Mutations in the *cueA* gene encoding a copper homeostasis P-type ATPase reduce the pathogenicity of *Pseudomonas aeruginosa* in mice. Int. J. Med. Microbiol. **295:**237-242.
221. **Scott JA.** 2007. The preventable burden of pneumococcal disease in the developing world. Vaccine **25:**2398-2405.
222. **Scott JE and Heatley F.** 1999. Hyaluronan forms specific stable tertiary structures in aqueous solution: a 13C NMR study. Proc. Natl. Acad. Sci. U. S. A. **96:**4850-4855.
223. **Shafeeq S, Kloosterman TG, Kuipers OP.** 2011. CelR-mediated activation of the cellobiose-utilization gene cluster in *Streptococcus pneumoniae*. Microbiology **157:**2854-2861.
224. **Shafeeq S, Kloosterman TG, Kuipers OP.** 2011. Transcriptional response of *Streptococcus pneumoniae* to $Zn^{(2+)}$ limitation and the repressor/activator function of AdcR. Metallomics. **3:**609-618.
225. **Shafeeq S, Yesilkaya H, Kloosterman TG, Narayanan G, Wandel M, Andrew PW, Kuipers OP, Morrissey JA.** 2011. The *cop* operon is required for copper homeostasis and contributes to virulence in *Streptococcus pneumoniae*. Mol. Microbiol. **81:**1255-1270.
226. **Shakhnovich EA, King SJ, Weiser JN.** 2002. Neuraminidase expressed by *Streptococcus pneumoniae* desialylates the lipopolysaccharide of Neisseria meningitidis and *Haemophilus influenzae*: a paradigm for interbacterial competition among pathogens of the human respiratory tract. Infect. Immun. **70:**7161-7164.
227. **Shelburne SA,3rd, Keith DB, Davenport MT, Horstmann N, Brennan RG, Musser JM.** 2008. Molecular characterization of group A *Streptococcus* maltodextrin catabolism and its role in pharyngitis. Mol. Microbiol. **69:**436-452.
228. **Shelburne SA, Davenport MT, Keith DB, Musser JM.** 2008. The role of complex carbohydrate catabolism in the pathogenesis of invasive streptococci. Trends Microbiol. **16:**318-325.

229. **SICARD AM.** 1964. A New Synthetic Medium for Diplococcus Pneumoniae, and its use for the Study of Reciprocal Transformations at the Amia Locus Genetics **50:**31-44.
230. **Sitthisak S, Knutsson L, Webb JW, Jayaswal RK.** 2007. Molecular characterization of the copper transport system in *Staphylococcus aureus*. Microbiology **153:**4274-4283.
231. **Smaldone GT and Helmann JD.** 2007. CsoR regulates the copper efflux operon *copZA* in *Bacillus subtilis*. Microbiology **153:**4123-4128.
232. **Solioz M, Abicht HK, Mermod M, Mancini S.** 2010. Response of gram-positive bacteria to copper stress. J. Biol. Inorg. Chem. **15:**3-14.
233. **Solioz M and Stoyanov JV.** 2003. Copper homeostasis in *Enterococcus hirae*. FEMS Microbiol. Rev. **27:**183-195.
234. **Song XM, Connor W, Hokamp K, Babiuk LA, Potter AA.** 2009. The growth phase-dependent regulation of the pilus locus genes by two-component system TCS08 in *Streptococcus pneumoniae*. Microb. Pathog. **46:**28-35.
235. **Spellerberg B, Rozdzinski E, Martin S, Weber-Heynemann J, Schnitzler N, Lutticken R, Podbielski A.** 1999. Lmb, a protein with similarities to the LraI adhesin family, mediates attachment of *Streptococcus agalactiae* to human laminin. Infect. Immun. **67:**871-878.
236. **Stroeher UH, Kidd SP, Stafford SL, Jennings MP, Paton JC, McEwan AG.** 2007. A pneumococcal MerR-like regulator and S-nitrosoglutathione reductase are required for systemic virulence. J. Infect. Dis. **196:**1820-1826.
237. **Tainer JA, Roberts VA, Getzoff ED.** 1991. Metal-binding sites in proteins. Curr. Opin. Biotechnol. **2:**582-591.
238. **Teitzel GM, Geddie A, De Long SK, Kirisits MJ, Whiteley M, Parsek MR.** 2006. Survival and growth in the presence of elevated copper: transcriptional profiling of copper-stressed *Pseudomonas aeruginosa*. J. Bacteriol. **188:**7242-7256.
239. **Tenenbaum T, Spellerberg B, Adam R, Vogel M, Kim KS, Schroten H.** 2007. *Streptococcus agalactiae* invasion of human brain microvascular endothelial cells is promoted by the laminin-binding protein Lmb. Microbes Infect. **9:**714-720.
240. **Terao Y, Kawabata S, Kunitomo E, Nakagawa I, Hamada S.** 2002. Novel laminin-binding protein of *Streptococcus pyogenes*, Lbp, is involved in adhesion to epithelial cells. Infect. Immun. **70:**993-997.
241. **Terra VS, Homer KA, Rao SG, Andrew PW, Yesilkaya H.** 2010. Characterization of novel beta-galactosidase activity that contributes to glycoprotein degradation and virulence in *Streptococcus pneumoniae*. Infect. Immun. **78:**348-357.
242. **Terzaghi BE and Sandine W.E.** 1975. Improved medium for lactic streptococci and their bacteriophages. Appl. Environ. Microbiol. **29:**807-813.
243. **Tettelin H, Nelson KE, Paulsen IT, Eisen JA, Read TD, Peterson S, Heidelberg J, DeBoy RT, Haft DH, Dodson RJ, Durkin AS, Gwinn M, Kolonay JF, Nelson WC, Peterson JD, Umayam LA, White O, Salzberg SL, Lewis MR, Radune D, Holtzapple E, Khouri H, Wolf AM, Utterback TR, Hansen CL, McDonald LA, Feldblyum TV, Angiuoli S, Dickinson T, Hickey EK, Holt IE, Loftus BJ, Yang F, Smith HO, Venter JC, Dougherty BA, Morrison DA, Hollingshead SK, Fraser CM.** 2001. Complete genome sequence of a virulent isolate of *Streptococcus pneumoniae*. Science **293:**498-506.
244. **Throup JP, Koretke KK, Bryant AP, Ingraham KA, Chalker AF, Ge Y, Marra A, Wallis NG, Brown JR, Holmes DJ, Rosenberg M, Burnham MK.** 2000. A genomic analysis of two-component signal transduction in *Streptococcus pneumoniae*. Mol. Microbiol. **35:**566-576.
245. **Titgemeyer F, Reizer J, Reizer A, Saier MH,Jr.** 1994. Evolutionary relationships between sugar kinases and transcriptional repressors in bacteria. Microbiology **140 (Pt 9):**2349-2354.
246. **Tomoyasu T, Tabata A, Imaki H, Tsuruno K, Miyazaki A, Sonomoto K, Whiley RA, Nagamune H.** 2012. Role of *Streptococcus intermedius* DnaK chaperone system in stress tolerance and pathogenicity. Cell Stress Chaperones **17:**41-55.
247. **Tong HH, Blue LE, James MA, DeMaria TF.** 2000. Evaluation of the virulence of a *Streptococcus pneumoniae* neuraminidase-deficient mutant in nasopharyngeal colonization and development of otitis media in the chinchilla model. Infect. Immun. **68:**921-924.
248. **Tong HH, James M, Grants I, Liu X, Shi G, DeMaria TF.** 2001. Comparison of structural changes of cell surface carbohydrates in the eustachian tube epithelium of chinchillas infected with a *Streptococcus pneumoniae* neuraminidase-deficient mutant or its isogenic parent strain. Microb. Pathog. **31:**309-317.
249. **Trappetti C, Kadioglu A, Carter M, Hayre J, Iannelli F, Pozzi G, Andrew PW, Oggioni MR.** 2009. Sialic acid: a preventable signal for pneumococcal biofilm formation, colonization, and invasion of the host. J. Infect. Dis. **199:**1497-1505.
250. **Tyx RE, Roche-Hakansson H, Hakansson AP.** 2011. Role of dihydrolipoamide dehydrogenase in regulation of raffinose transport in *Streptococcus pneumoniae*. J. Bacteriol. **193:**3512-3524.

251. **Ulijasz AT, Andes DR, Glasner JD, Weisblum B.** 2004. Regulation of iron transport in *Streptococcus pneumoniae* by RitR, an orphan response regulator. J. Bacteriol. **186:**8123-8136.
252. **Ulijasz AT, Falk SP, Weisblum B.** 2009. Phosphorylation of the RitR DNA-binding domain by a Ser-Thr phosphokinase: implications for global gene regulation in the streptococci. Mol. Microbiol. **71:**382-390.
253. **van Hijum SAFT, Garcia de la Nava J, Trelles O, Kok J, Kuipers OP.** 2003. MicroPreP: a cDNA microarray data pre-processing framework. Appl. Bioinformatics. **2:**241-244.
254. **van Hijum SAFT, de Jong A, Baerends RJS, Karsens HA, Kramer NE, Larsen R, den Hengst CD, Albers CJ, Kok J, Kuipers OP.** 2005. A generally applicable validation scheme for the assessment of factors involved in reproducibility and quality of DNA-microarray data. BMC Genomics **6:**77.
255. **van Opijnen T, Bodi KL, Camilli A.** 2009. Tn-seq: high-throughput parallel sequencing for fitness and genetic interaction studies in microorganisms. Nat. Methods **6:**767-772.
256. **van TH and Declerck N.** 2001. Structural insights into the regulation of bacterial signalling proteins containing PRDs. Curr. Opin. Struct. Biol. **11:**685-693.
257. **Veldhuis NA, Gaeth AP, Pearson RB, Gabriel K, Camakaris J.** 2009. The multi-layered regulation of copper translocating P-type ATPases. Biometals **22:**177-190.
258. **Versieck J.** 1985. Trace elements in human body fluids and tissues. Crit. Rev. Clin. Lab. Sci. **22:**97-184.
259. **Wakeman CA and Skaar EP.** 2012. Metalloregulation of Gram-positive pathogen physiology. Curr. Opin. Microbiol. **15(2):**169-174.
260. **Waldron KJ, Rutherford JC, Ford D, Robinson NJ.** 2009. Metalloproteins and metal sensing. Nature **460:**823-830.
261. **Ward SK, Hoye EA, Talaat AM.** 2008. The global responses of *Mycobacterium tuberculosis* to physiological levels of copper. J. Bacteriol. **190:**2939-2946.
262. **Watson DA and Musher DM.** 1990. Interruption of capsule production in *Streptococcus pneumoniae* serotype 3 by insertion of transposon Tn916. Infect. Immun. **58:**3135-3138.
263. **Webb AJ, Homer KA, Hosie AH.** 2007. A phosphoenolpyruvate-dependent phosphotransferase system is the principal maltose transporter in *Streptococcus mutans*. J. Bacteriol. **189:**3322-3327.
264. **Weiser JN, Austrian R, Sreenivasan PK, Masure HR.** 1994. Phase variation in pneumococcal opacity: relationship between colonial morphology and nasopharyngeal colonization. Infect. Immun. **62:**2582-2589.
265. **Weston BF, Brenot A, Caparon MG.** 2009. The metal homeostasis protein, Lsp, of *Streptococcus pyogenes* is necessary for acquisition of zinc and virulence. Infect. Immun. **77:**2840-2848.
266. **Whitman WB, Coleman DC, Wiebe WJ.** 1998. Prokaryotes: the unseen majority. Proc. Natl. Acad. Sci. U. S. A. **95:**6578-6583.
267. **Wizemann TM, Heinrichs JH, Adamou JE, Erwin AL, Kunsch C, Choi GH, Barash SC, Rosen CA, Masure HR, Tuomanen E, Gayle A, Brewah YA, Walsh W, Barren P, Lathigra R, Hanson M, Langermann S, Johnson S, Koenig S.** 2001. Use of a whole genome approach to identify vaccine molecules affording protection against *Streptococcus pneumoniae* infection. Infect. Immun. **69:**1593-1598.
268. **Wolschendorf F, Ackart D, Shrestha TB, Hascall-Dove L, Nolan S, Lamichhane G, Wang Y, Bossmann SH, Basaraba RJ, Niederweis M.** 2011. Copper resistance is essential for virulence of *Mycobacterium tuberculosis*. Proc. Natl. Acad. Sci. U. S. A. **108:**1621-1626.
269. **Wood DM, Brennan AL, Philips BJ, Baker EH.** 2004. Effect of hyperglycaemia on glucose concentration of human nasal secretions Clin. Sci. (Lond) **106:**527-533.
270. **Woods DR and Reid SJ.** 1995. Regulation of nitrogen metabolism, starch utilisation and the beta-hbd-adh1 gene cluster in *Clostridium acetobutylicum*. FEMS Microbiol. Rev. **17:**299-306.
271. **Yesilkaya H, Kadioglu A, Gingles N, Alexander JE, Mitchell TJ, Andrew PW.** 2000. Role of manganese-containing superoxide dismutase in oxidative stress and virulence of *Streptococcus pneumoniae*. Infect. Immun. **68:**2819-2826.
272. **Yesilkaya H, Manco S, Kadioglu A, Terra VS, Andrew PW.** 2008. The ability to utilize mucin affects the regulation of virulence gene expression in *Streptococcus pneumoniae*. FEMS Microbiol. Lett. **278:**231-235.
273. **Yesilkaya H, Soma-Haddrick S, Crennell SJ, Andrew PW.** 2006. Identification of amino acids essential for catalytic activity of pneumococcal neuraminidase A. Res. Microbiol. **157:**569-574.
274. **Yesilkaya H, Spissu F, Carvalho SM, Terra VS, Homer KA, Benisty R, Porat N, Neves AR, Andrew PW.** 2009. Pyruvate formate lyase is required for pneumococcal fermentative metabolism and virulence. Infect. Immun. **77:**5418-5427.
275. **Zeng L and Burne RA.** 2009. Transcriptional regulation of the cellobiose operon of *Streptococcus mutans*. J. Bacteriol. **191:**2153-2162.
276. **Zeng L and Burne RA.** 2008. Multiple sugar: phosphotransferase system permeases participate in catabolite modification of gene expression in *Streptococcus mutans*. Mol. Microbiol. **70:**197-208.
277. **Zeng L, Das S, Burne RA.** 2010. Utilization of lactose and galactose by *Streptococcus mutans*: transport, toxicity, and carbon catabolite repression. J. Bacteriol. **192:**2434-2444.

278. **Zeng L, Martino NC, Burne RA.** 2012. Two gene clusters coordinate galactose and lactose metabolism in *Streptococcus gordonii*. Appl. Environ. Microbiol. **78(16):**5597-5605.
279. **Zhang Q, Choo S, Finn A.** 2002. Immune responses to novel pneumococcal proteins pneumolysin, PspA, PsaA, and CbpA in adenoidal B cells from children. Infect. Immun. **70:**5363-5369.
280. **Zhang Z, Aboulwafa M, Smith MH, Saier MH,Jr.** 2003. The ascorbate transporter of *Escherichia coli*. J. Bacteriol. **185:**2243-2250.
281. **Zomer AL, Buist G, Larsen R, Kok J, Kuipers OP.** 2007. Time-resolved determination of the CcpA regulon of *Lactococcus lactis* subsp. cremoris MG1363. J. Bacteriol. **189:**1366-1381.

Nederlandse samenvatting
(Dutch Summary)

Streptococcus pneumoniae is een grampositieve menselijke ziekteverwekker die jaarlijks verantwoordelijk is voor de dood van miljoenen mensen, voornamelijk kinderen, en veroorzaakt pneumokokken infecties zoals longontsteking, hersenvliesontsteking, middenoor ontsteking en sepsis. De bacterie leeft onder normale omstandigheden in de neusholte. Soms kan het zich, als de omstandigheden juist zijn, door het lichaam naar andere plekken verplaatsen en daar infecties veroorzaken. Om in deze zeer verschillende gastheer omgevingen te overleven en infecties te veroorzaken moet *S. pneumoniae* deze aanzienlijke variaties in omgevingsomstandigheden kunnen registreren en zich daaraan aan kunnen passen. Hieronder vallen de veranderingen in concentraties metaalionen in gastheer weefsel en de beschikbaarheid van verschillende koolstofbronnen. Er is weinig bekend over het mechanisme van interactie van deze humane pathogeen met de veranderende omgeving in het menselijk lichaam. Dit proefschrift beschrijft de gen gereguleerde respons van *S. pneumoniae* op de variërende metaal en koolhydraat beschikbaarheid. Wisselende metaal concentraties en verschillende koolhydraat bronnen zijn factoren waar hij mee geconfronteerd kan worden in verschillende plekken van het menselijk lichaam tijdens de infectie. Een aantal regulatoren die reageren op verschillende metaal (b.v. zink en koper) en koolhydraat bronnen (b.v. cellobiose) zijn gekarakteriseerd en hun regulonen zijn bestudeerd. De resultaten beschreven in dit proefschrift zullen helpen bij het beter begrijpen van de moleculaire biologie van *S. pneumoniae* en verschaffen meer inzicht in de respons van *S. pneumoniae* op veranderende omgevingscondities.

Zink is een belangrijk, maar toxisch, zwaar metaal ion dat de expressie van een aantal (virulentie) genen in streptococci reguleert. Eerder hebben we de genoom-brede respons van S. pneumoniae op zink-stress geanalyseerd. In **hoofdstuk 2** hebben we een analyse van het transcriptoom gedaan om genen te identificeren die differentieel tot expressie komen bij intracellulaire zink limitatie. Dit leidde tot een aantal genen die sterk omhoog gereguleerd zijn in de afwezigheid van extracellulaire zink. Hieronder vallen de genen die behoren tot het regulon van de zink-responsieve repressor AdcR, zoals *adcBCA*, coderend voor een zink-afhankelijk ABC-opname systeem, *adcAII*, coderend voor een zink-bindend lipoproteïne, en ook virulentie genen behorend tot de Pht familie (PhtA, PhtB, PhtD en PhtE). Met transcriptoom analyse, *lacZ*-reporter studies, *in vitro* DNA-binding experimenten, en *in silico* operator voorspellingen, laten we zien dat AdcR direct de promoters van *adcRCBA*, *adcAII-phtD*, *phtA*, *phtB* en *phtE* onderdrukt in de aanwezigheid van zink. AdcR kan ook

functioneren als een activator, omdat het in de aanwezigheid van zink direct de expressie van *adh*, coderend voor een zink-bevattend alcohol dehydrogenase, induceert.

Net als zink, zijn hoge concentraties koper ook giftig voor bacteriën. Om schade aan de cel te voorkomen moet het aantal koperionen laag zijn en blijven. In **hoofdstuk 3** beschrijven we dat de expressie van een aantal pneumococcus genen gereguleerd worden door koper zoals; een operon dat met een gen dat de CopY transcriptionele regulator codeert, een *cupA* gen dat een eiwit met onbekende functie codeert en het *copA* gen voor P1-type ATPase. Transcriptie analyse laat zien dat het *cop*-operon *in vitro* wordt geïnduceerd terwijl het door toevoeging van zink gerepresseerd wordt. Verder vindt autoregulatie plaats via de CopY repressor. We laten ook zien dat de CopA ATPase de belangrijkste rol vervult in koperresistentie en bewijzen dat naast CopA ook CupA een rol speelt in deze resistentie. Onze resultaten laten zien dat koper homeostatis belangrijk is voor pneumococcus virulentie omdat de expressie van het *cop*-operon geïnduceerd wordt in de longen en in de nasopharynx van intranasaal geïnfecteerde muizen. Ook zien we een gereduceerde virulentie in een muismodel van de *copA* mutant, die *in vitro* een verminderede groei laat zien bij hoge koper concentraties. Tenslotte laten we voor het eerst zien dat koper homeostasis voor pneumokokken nodig blijkt te zijn om te overleven in de nasopharynx.

Hoofdstukken 4 en **5** beschrijven het effect dat het aanbieden van cellobiose heeft op genexpressie in *Streptococcus pneumoniae* alsook de functionele karakterisatie van de regulatoren CelR en CelR-II die op cellobiose reageren. Het humane pathogeen *S. pneumoniae* bevat veel genen die coderen voor fosfotransferrasesystemen (PTS) en ABC (ATP-bindende cassette) suikertransporters, waaronder systemen voor het opnemen van het β-glucosidasesuiker cellobiose. In deze studie tonen we aan dat regulator van transcriptie, CelR, waarvan voorheen bewezen is dat het een belangrijke rol speelt in pneumococcen virulentie, expressie stimuleert van het gencluster dat codeert voor eiwitten die nodig zijn voor het kunnen gebruiken van cellobiose door *S. pneumoniae* (*cel* locus). Genexpressie van de *cel* locus, aangestuurd door 2 promoters, was verhoogd wanneer cellobiose de enige koolstofbron in het medium was, terwijl genexpressie omlaag gaat bij aanwezigheid van glucose in het medium. Ook hebben we een sequentie van 22-bp geïdentificeerd in de promoters van *celA* en *celB* (5'-YTTTCCWTAWCAWTWAGGAAAA-3') dat vermoedelijk een herkenningssequentie is voor de regulator CelR. *In silico* analyse toonde tevens aan dat deze sequentie zeer geconserveerd is onder pathogene streptococci. Ingekorte promoters van *celA* en *celB*, waarbij de helft van de CelR regulator sequentie verwijderd werd, bevestigden dat de CelR bindingssequentie in promoters P*celA* en P*celB* functioneel is. Uitkomsten van

transcriptoom studies van een *celR* mutant in combinatie met *in silico* voorspellingen van de CelR bindingssequentie in het volledige D39 genoom sequentie tonen aan dat het *cel* locus het enige cluster is waarvan de genen onder controle staan van CelR. Daarom concluderen we dat CelR een specifieke regulator voor cellobiose-afhankelijke transcriptie van *cel* locus genen.

Voor **hoofdstuk 5,** hebben we de regulatie van transcriptie van het *cel-II* operon gekarakteriseerd. In tegenstelling tot de *cel* locus, is het *cel-II* operon zeer geconserveerd in alle gesequenste *S. pneumoniae* stammen wat een aanwijzing is voor diens belangrijke rol in de levenscyclus pneumococci. Expressie van het *cel-II* operon is verhoogd in aanwezigheid van cellobiose, terwijl het aanbieden van glucose de expressie van het operon verlaagt. Van een nieuw gevonden GntR-achtige transcriptie regulator (CelR-II) is aangetoond dat deze functioneert als repressor van transcriptie van het *cel-II* operon en dat de repressie werd opgeheven door de aanwezigheid van cellobiose. Van CelR-II repressie werd ook aangetoond dat dit tot stand kwam via een 20-bp DNA operator sequentie (5'-AAAAATGTCTAGACAAATTT-3') die aanwezig is in promoter P*celA-II* zoals bevestigd middels experimenten met ingekorte promoters.

In **hoofdstuk 6** zijn de regulatoire mechanismen van het gencluster *SPD0424-428*, mogelijk coderend voor een cellobiose/lactose-specifieke PTS, onderzocht. Wij demonstreren dat dit gencluster wordt afgeschreven als een transcriptionele unit afkomstig van de promoter van het gen *SPD0424*. Upstream van *SPD0424* is een gen geïdentificeerd dat codeert voor een ROK familie geclassificeerde transcriptionele regulator (RokA: SPD0423). DNA microarray- en transcriptionele reporter analyse met een *rokA* mutant, tonen aan dat RokA werkt als een transcriptionele repressor op het *SPD0424-28* operon. Daarnaast is er een 25-bp DNA operator site (5'-TATATTTAATTTATAAAAAATAAAA-3') in de promoter regio van *SPD0424* geïdentificeerd, die gevalideerd is door middel van promoter mutatie studies, DNaseI footprinting en elektroforese-mobiliteit-verschuiving-assays (EMSAs). De invloed van verschillende suikerbronnen op de expressie van het operon is onderzocht. Hierbij zijn alleen kleine variaties in expressie van het operon waargenomen. Hierdoor kan er geen inducerend substraat worden aangewezen voor de transcriptionele regulatie van RokA.

Concluderende opmerkingen

Uit voorgaande studies is gebleken dat, indien men de infecties door pneumococcen wil begrijpen, informatie nodig is van de omgevingscondities die *S. pneumoniae* cellen trotseren in de gastheer. Dit proefschrift biedt inzicht in het effect van omgevingsstimuli,

zoals verschillende concentraties metaalionen en de beschikbaarheid van verschillende koolwaterstofverbindingen, op de genexpressie van de humane pathogeen *S. pneumoniae*. De genen van dit organisme reageren op veranderde concentraties metalen en aanwezige koolwaterstofverbindingen, en zijn hierbij eveneens gekoppeld aan de virulentie van *S. pneumoniae*. Aldus kan met het begrijpen van de regulatie van deze genen een stap voorwaarts worden gezet naar de ontwikkeling van een methode om infecties van *S. pneumoniae* te voorkomen.

Acknowledgements

The acknowledgement is always a difficult part to write because most of the people start reading it before the thesis itself. First of all I would like to acknowledge the blessing of God that leads me on every path of my life and has brought success in my life.

Foremost, I would like to express my sincere gratitude to my promoter Oscar Kuipers for giving me an opportunity to work in his group, for continuously supporting my PhD research, for the freedom he let me in choosing the different project directions, for his trust, patience, motivation, enthusiasm, guidance and immense knowledge, for his encouragement to present research work in international conferences and for much more. His guidance helped me all the time: in research as well as in the writing of this thesis. I could not have imagined having a better mentor for my PhD study.

I am extremely thankful to my co-promoter, a pianist and a good friend, Tomas Kloosterman for his moral support, motivation and advises, for his continuous support in my PhD research, for his daily help in the lab, for teaching me his top laboratory secrets and for much more. Thanks for contaminating my soul with classical music. I am really glad that we have worked together.

My sincere thanks go to Jan-Willem Veening for giving me a Postdoc opportunity in his team on exciting project.

My sincere gratitude goes to Jan Kok for all the scientific and social discussions and immense advices during the Molgen work discussions.

I would like to acknowledge Higher Education Commission (HEC) of Pakistan for their financial support. Further, I would like to thank Jan-Maarten van Dijl, J. D. van Elsas and Peter. Hermans for agreeing on being a part of the reading committee, reading this book and giving their approval.

I would like to thank my collaborators Hasan Yesilkaya, Julie Morrissey Geetha Narayanan and Peter Andrew.

I would like to thank all the present and former Molgen members for their supports and ideas: Agata, Akos, Aleksandra, Andrius, Anita, Anja, Araz, Anne-Rute, Anne H, Anne de J, Anne S, Ana S, Auke, Bogusia, Claire, Dongdong, Evert-Jan, Ganesh, Hein, Harma, Imke, Jetta, Jelle, João, Jannet, Joanna (Asia), Julio, Jeroen, Katrin, Laetitia, Lidia, Liang, Lieke, Maarten, Marta, Mozes, Martijn, Mirjam, Maria, Marielle, Mirelle, Manolo, Michel, Morten, Pamela, Peter, Robin, Robyn, Rustum, Rutger, Ruud, Sandra, Sjoerd, Tonia, Tomas, Tom, Vijay, Wout and Yoshimitsu. You all gave me such a pleasant and unforgettable time in Molgen.

I would like to thank my grandparents and parents who always encouraged me for higher education. Without their motivation and prayers, it would not have been possible for me to reach this stage of life. I would like to thank my younger sister, a little angel for helping me with computers and for editing my thesis. I would like to thank my brother and my best childhood friend for his motivation, scientific advices and for reading my thesis.

I am really thankful to all my Pakistani friends and colleagues for making me feel like a home in Groningen: Ishtiaq, Atta, Javaid, Ali, Afzal, Irfan, Saleem, Kokab, Nadia, Tanveer, Waqas K, Ijaz, Waqas A, Shafique, Tariq, Afzal S. Omer, Zia, Umer, Adnan, Latif, Munir, Irfan A, Bailal, Samin, Ehsan, Rashid, Aisha, and Irshad. Special thanks go to the families of Ali, Ishtiaq and Tanveer for inviting me to the celebrations of all Pakistani events. I am really thankful to Ali, Afzal, Irfan and Tariq for the lunch discussions and cooking delicious food for me. Atta, I am really thankful that you involved me in social activities to relieve the scientific stress. Javaid thanks for playing table tennis and chess during the evening pauses of long days in the lab.

Cricket is a game of passion and as important as religion in South Asia. I would like to thank all the members of GCC (Groninger Cricket Club) especially Sarfraz jee, Yusuf, Shariq, Sanger, Raj, Adriaan, Gerard, Gopi, Savneet, Suleman, Shiva, Soumen, Sunil, Naveen and Yash. I wish a lot of success to the GCC in the coming years.

I would like to thank all the persons that helped me in any respect for giving a shape to this book.

List of Publications

In Preparation:

1. **Shafeeq. S**, O.P. Kuipers and T.G. Kloosterman. 2012. Cellobiose-mediated gene expression in *Streptococcus pneumoniae* and the repressor function of the novel GntR-type regulator CelR-II.

Submitted:

2. **Shafeeq. S**, T.G. Kloosterman, V. Rajendran and O.P. Kuipers. 2012. Characterization of the ROK family transcriptional regulator RokA of *Streptococcus pneumoniae* D39.
3. **Shafeeq. S**, O.P. Kuipers and T.G. Kloosterman. 2012. The role of zinc and related metal ions in the interplay between pathogenic streptococci and their hosts.

Published:

4. **Shafeeq. S**, T.G. Kloosterman and O.P. Kuipers. **2011**. CelR-mediated activation of the cellobiose-utilization locus in *Streptococcus pneumoniae*. **Microbiology.** 157(Pt 10): 2854-61.
5. **Shafeeq. S**, H. Yesilkaya, T.G. Kloosterman, G. Narayanan, M. Wandal, P.W. Andrew, O.P. Kuipers, and J. A. Morrissey. **2011**. The *cop* operon is required for copper homeostasis and virulence in *Streptococcus pneumoniae***. Molecular Microbiology.** 81 (5): 1255-70.
6. **Shafeeq. S**, T. G. Kloosterman and O.P. Kuipers. **2011**. Transcriptional response of *Streptococcus pneumoniae* to Zn^{2+} limitation and the repressor/activator function of AdcR. **Metallomics.** 3(6): 609-18.
7. **Shafeeq, S.**, M. Rahman, and Y. Zafar. **2006**. Genetic variability of different wheat (*Triticum aestivum* L.) genotypes/cultivars under induced water stress. **Pakistan Journal of Botany.** 38(5): 1671-78.